



通信网络精品图书

视频会议系统 原理与测试

王湘宁 顾晓鹏 陈华磊 卢洪勇 编著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY



通信网络精品图书

视频会议系统 原理与测试

- ✦ 视频会议跨界电信领域与广电领域，不当当作电信宽带业务的一种
- ✦ 运用视频图像分析技术，开辟视频会议系统测试的新领域



策划编辑：宋 梅
责任编辑：宋 梅
封面设计：张 昱

ISBN 978-7-121-21911-5



定价：39.00 元

通信网络精品图书

视频会议系统 原理与测试

王湘宁 顾晓鹏 陈华磊 卢洪勇 编著

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书主要描述视频会议系统的原理及其测试技术,全书共5章,第1章主要介绍视频会议的发展历程、主要功能和典型构成;第2章描述光度学和色度学的相关知识;第3章介绍现代电视技术基础,第2、3章的内容是视频技术的理论和技术基础;第4章重点介绍视频网络技术,重点对H.323会议技术、SIP会议技术和基于IMS的会议技术所涉及的协议流程进行介绍;第5章阐述视频会议系统测试技术,对视频会议系统功能测试、单节点设备性能测试、视频图像质量评测、视频网络QoS指标测试进行详细的阐述。

本书适合从事电信视频业务相关运营和测试等工作的工程技术人员阅读,也可作为相关高等院校师生的参考书。

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

视频会议系统原理与测试 / 王湘宁等编著. —北京:电子工业出版社, 2014.1

(通信网络精品图书)

ISBN 978-7-121-21911-5

I. ①视… II. ①王… III. ①电视会议系统 IV. ①TN948.63

中国版本图书馆CIP数据核字(2013)第274721号

策划编辑:宋 梅

责任编辑:宋 梅

印 刷:三河市鑫金马印装有限公司

装 订:三河市鑫金马印装有限公司

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:12 字数:307千字

印 次:2014年1月第1次印刷

印 数:3000册 定价:39.00元

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 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

序

视频会议业务作为一项基础电信业务在电信公网和行业专网得到广泛的部署和应用，视频会议业务实现了人们远距离“临场感”彼此交流的愿望，在提高交流效率的同时也极大降低了交流成本。

视频会议系统的技术和相应的标准一直在不断发展中，总体来说主要经历了模拟视频到数字视频、电路交换到 IP 分组交换、标清视频到高清视频、单一业务到融合业务等发展过程。视频会议技术属于广电领域电视技术与电信领域通信技术相互交叉的学科，随着高清电视技术和宽带通信技术的不断发展，融合应用上述两个领域发展成果的视频会议系统也在不断推陈出新。一方面，视频会议系统的图像分辨率会越来越高，会议的音、视临场感会越来越强；另一方面，用户的接入手段也会越来越丰富，从专业的视频会议终端到手持移动通信终端均可等同无缝地加入视频会议系统中。这样一来，无论人们身在何处，均可方便地得到相应的视频会议服务。

本书写作团队多年来承担中国联通“宝视通”视频会议业务研究和测试，在研究和测试过程中，发现视频会议系统缺乏较为完整的测试体系，通常只将视频会议业务当作一项普通的数据业务。为此，该技术团队通过对测试实践工作的总结，希望能将图像质量评价方法与视频传输技术对其影响有机结合起来，从而形成更为系统的视频会议系统测试评价体系和方法。

本书的内容包括视频会议基础和发展演进、视频会议的多媒体技术基础和电视技术基础、视频会议主要协议框架和视频会议系统测试技术。本书的特色是在介绍视频会议技术基本原理和协议框架的基础上，对视频会议测试方法进行系统性阐述和分析。希望本书的出版，能给从事视频会议系统测试和研究的工程技术人员提供有益的参考，也为视频会议系统测试技术的进一步研究奠定良好的基础。

唐雄燕

中国联通网络技术研究院首席专家

2013 年 11 月于北京

前 言

视频会议，是指处于异地的多个个体或群体，通过通信网络和视频设备，彼此实时地交换动态视频和语音，同时辅之以静态图像、文字等材料的交流，从而实现不用身处一室便可实现千里之外的“聚聚一堂”。

视频会议从产生之初至今，正经历着网络技术、用户接入方式、用户体验和使用功能的全方位发展。从网络技术发展的角度来看，视频会议经历了从两点通信到大规模多点通信，从电路交换到包交换的过程，目前正向 IMS 网络融合的方式演变。从用户的接入方式看，视频会议用户接入从简单的电路接入发展到无线、以太网、光纤等多种接入手段。从实际的用户体验来看，已经从模拟电视技术发展到了数字电视技术，从标清视频发展到高清视频。特别是近年来发展起来的高清“网真”会议系统，几乎可以使电子会场“乱真”，与会者往往忘记自己是处在“虚拟”的会场中。从视频会议的使用功能来看，从简单的视频和音频信号交换的单一功能，发展到能模拟实际会场并全面支持用户实际交流所需的各项功能。

视频会议作为电信公网与行业专网的一项重要业务，正随着高清电视技术和宽带通信技术的发展而得到广泛的推广应用。因此，在国家大力推进节能减排和可持续发展的今天，推广视频会议，无疑是顺应时代发展潮流、节约成本、提高效率的重要举措。

视频会议技术是集电信网络通信技术与电视视频技术相融合的交叉学科，同时涉及电信领域和广电领域相关标准。目前，电信运营商视频会议业务处于重要发展期，但缺乏视频会议系统的完整测试体系，内容指标测试也并不十分深入。希望本书的出版，能较为系统地介绍视频会议所涉及的图像视频技术和测试评估方法，为广大从事视频业务的电信运营工程技术人员在测试时提供一定的帮助和参考，从此摒弃把视频会议业务简单地作为一种电信数据业务的笼统思维方式。

本书共 5 章，第 1 章主要介绍视频会议的发展历程、主要功能和典型构成；第 2 章描述光度学和色度学的相关知识；第 3 章介绍现代电视技术基础，第 2、3 章的内容是视频技术的理论和技术基础；第 4 章重点介绍视频网络技术，重点对 H.323 会议技术、SIP 会议技术和基于 IMS 的会议技术所涉及的协议流程进行介绍；第 5 章阐述视频会议系统测试技术，对视频会议系统功能测试、单节点设备性能测试、视频图像质量评测、视频网络 QoS 指标测试进行详细的阐述。其中，第 1~3 章和第 5 章的前 3 节由王湘宁编写，第 4 章 4.1 节由顾晓鹏编写，第 4 章 4.2、4.3 节由

陈华磊编写，第 5 章 5.4 节由卢洪勇编写。全书由王湘宁统稿。宋雪娜、多红月也参与了本书的编写工作，提供了相关技术资料，主要负责图表制作工作。

本书的编撰以中国联通“宝视通”视频会议测试项目为基础，在编写的过程中得到了华为公司、Polycom 公司等相关视频会议设备厂商的协助，在此向其表示诚挚的谢意！

鉴于作者水平，书中难免会存在片面甚至错误之处，还望业界及相关领域专家予以批评指正！

编著者

2013 年 11 月于北京

目 录

第 1 章	视频会议系统概述	1
1.1	视频会议系统的发展历程与趋势	1
1.2	视频会议系统的主要功能	8
1.3	视频会议系统设计方案	11
1.3.1	视频会议系统的设计原则	11
1.3.2	华为公司视频会议系统解决方案	12
第 2 章	光度与色度	19
2.1	光度学知识	19
2.1.1	可见光	19
2.1.2	常用的光学物理量	19
2.2	人眼的视觉特性	22
2.2.1	人眼对光亮度的感觉	22
2.2.2	人眼对光色度的感觉	23
2.2.3	人眼的分辨力	26
2.2.4	人眼视觉惰性	28
2.2.5	人眼的视觉适应性与图像对比度	29
2.3	色度学知识	30
2.3.1	彩色三要素	30
2.3.2	1931 CIE-RGB 计色系统	30
2.3.3	1931 CIE-XYZ 计色系统	33
2.3.4	1960 CIE-UCS 计色系统	36
2.3.5	1964 CIE-UCS 均匀颜色空间 (W^* 、 U^* 、 V^*)	39
2.3.6	1976 CIE- $L^* u^* v^*$ 均匀颜色空间	39
第 3 章	现代电视技术基础	41
3.1	模拟电视技术基础	41
3.1.1	逐行扫描与隔行扫描	41
3.1.2	亮度方程式	44
3.1.3	NTSC 制电视信号	45
3.1.4	PAL 制电视信号	47
3.2	数字电视技术基础	48
3.2.1	视频信号采样格式	49
3.2.2	视频信号的参数	50
3.2.3	视频压缩编码技术	53

3.2.4	视频压缩编码标准	57
3.2.5	系统复用	61
3.3	主要视频接口特性	62
第 4 章	视频会议网络技术	65
4.1	H.323 视频会议网络技术	66
4.1.1	H.323 协议栈	69
4.1.2	RAS 信令协议	70
4.1.3	H.225.0 呼叫信令协议	75
4.1.4	H.245 控制信令协议	79
4.1.5	H.323 呼叫流程实例	86
4.2	基于 SIP 协议的视频会议网络技术	87
4.2.1	SIP 的逻辑实体	88
4.2.2	SIP 基本消息	90
4.2.3	SDP 协议概述	92
4.2.4	实时传输协议 (RTP)	100
4.2.5	实时传输控制协议 (RTCP)	101
4.2.6	基于 SIP 协议的视频会议系统	111
4.3	基于 IMS 技术的会议	121
4.3.1	IMS 的概述	121
4.3.2	IMS 会议结构	124
4.3.3	控制信令 SIP 概述	126
4.3.4	媒体信令 H.248 概述	126
4.3.5	IMS 的会议系统特点	127
4.3.6	会议系统需求分析	128
4.3.7	会议系统的功能	129
4.3.8	会议系统在 IMS 中的位置	130
4.3.9	信令流程	134
第 5 章	视频会议系统的测试方法	143
5.1	视频会议系统功能测试研究	143
5.1.1	基本业务能力测试	144
5.1.2	兼容与互通性测试	147
5.2	视频会议图像质量测试研究	148
5.2.1	视频图像质量主观测试	149
5.2.2	视频图像质量客观测试	152
5.3	视频会议显示器性能测试研究	156
5.3.1	亮度测试	157
5.3.2	对比度测试	159
5.3.3	色域覆盖率测试	160

5.3.4	清晰度测试	160
5.4	视频会议网络 QoS 性能测试研究	161
5.4.1	网络时延的基本概念和时延测量原理	162
5.4.2	时延测量中的关键技术和影响因素	163
5.4.3	网络时延测量方法	165
5.4.4	抖动测试方法	166
5.4.5	Smartbits 丢包率测量方法	170
5.4.6	利用 Ping 命令进行快速网络性能测试	174
缩略语		177
参考文献		182

第 1 章 视频会议系统概述

随着 IP 网络及通信技术的飞速发展，传统的语音和文字通信方式已不能满足人们日益丰富的现代生活需要，人们迫切希望能有一种音频、视频、数据于一体的多媒体互动交流方式，多媒体通信技术正是在这种背景下发展起来的一种新技术。通过多媒体通信技术，我们可以不受时空的限制来进行集图、声、像、文等多种媒体形式为一体的多媒体信息交流，通信双方不但可以相互听到对方的声音，还可以传送数据文件甚至进行面对面的视频会话。因此，多媒体通信技术的出现与发展使得视频图像的网络传输成为可能。

视频会议系统，又称为视讯会议系统，是一种能把声音、图像、文本等多种信息从一个地方传送到另一个地方的通信系统。采用视频会议的方式，可以使身处两个会场或多个会场的与会者，既能听到其他与会人员的声音，又能看到其他会场的图像，增强临场感，同时还可以通过传送文件、共享电子白板等方式讨论相关问题等，使与会者如同身临其境在现场开会一样，从而提高工作效率，降低远距离会议费用。

视频会议系统的技术和相应的标准一直处于不断的发展过程之中，总的来说，主要经历了模拟视频到数字视频、电路交换通信到 IP 分组通信、标清图像到高清图像、单一业务到融合业务等过程。

1.1 视频会议系统的发展历程与趋势

1. 视频传输技术的发展历程

20 世纪 60 年代发达国家就开始了模拟视频会议系统的研究，并逐渐商用化。早期的视频会议系统以模拟方式传输，尽管传送的只是黑白图像，并只局限于在两个地点之间，但视频会议还是要占用很大的带宽，费用很高，因此没有得到发展。其代表有最早研制出来的美国贝尔实验室的可视电话、英国 BT 公司的 1 MHz 带宽的黑白会议电视系统。

20 世纪 80 年代以来，视频会议在理论研究和实用系统研制方面得到迅速发展，图像编解码技术取得突破，信道费用降低，为视频会议走向大规模应用提供

了很好的发展条件。这个时期，数字视频会议系统取代了模拟视频会议系统，在某些地区开始形成视频会议网，如 20 世纪 80 年代初期，日本和美国研制出的 2 Mbps 彩色数字视频会议系统，形成非标准的国内视频会议网。但这一时期的数字视频会议系统一直使用专用的编解码硬件和软件，要求互通的会议终端使用的编解码器必须来自同一个厂商，否则不能正常工作，从而极大地阻碍了视频会议系统的可扩充性和各系统之间的互操作性。

1984 年，ITU-T 组织制定出适合 E1/T1 通信网的 H.100 系列建议，首次为可视电话和视频会议制定了标准。1988 年，ITU-T 通过了 H.261 标准，该标准不仅规定了编解码器的主要结构，也规定了视频会议视频信号的组成、纠错方法和数据结构，为各种产品的国际间互通提供了保障，从而出现了国际统一标准的视频会议系统，极大地推动了视频会议的发展。为了便于各个电视会议系统设备供应商的产品实现互通，1990 年，ITU-T 又制定了 H.320 系列标准，对电视会议系统的性能指标、压缩算法、信息结构、控制命令、规程和组建电视会议网的原则作了完整的规定，促进了电视会议的健康发展。

1990 年至 1995 年是 ISDN 网络上数字视频会议的发展阶段，该阶段是视频会议技术发展较快的一个时期，这主要得益于数字压缩技术的进步（例如，ITU-T 通过了 H.261 建议）以及 ISDN 技术和标准的成熟。在此阶段，一些视频会议厂家生产出多种基于 H.320 的彼此互相兼容的视频会议设备，世界各国也建设了不同规模的视频会议网络。

20 世纪 90 年代后期，基于 TCP/IP 协议的 Internet 网络规模、用户数量以及业务量呈指数型增长，对电信业产生了巨大而深刻的影响，进而基于分组交换网络的多媒体通信系统逐渐成为研究人员和电信、网络及计算机厂商关注的焦点。随着 ITU-T 建议 H.323 的出台，基于 IP 网络的视频会议技术的应用也相应展开，视频会议系统得到空前地发展，在政府、军事、金融、电信、教育、企业等领域的应用也越来越广泛。

迄今为止，ITU-T 组织已发展了在各种网络上传输电视会议信号的系列建议等。其中，H.320 是在窄带综合业务数字网（N-ISDN）上进行多媒体通信的标准，H.321 是在宽带综合业务数字网（B-ISDN）上进行多媒体通信的标准，H.322 是在有服务质量（QoS）确保的局域网（LAN）上进行多媒体通信的标准，H.323 是在无服务质量（QoS）确保的分组网络（Packet-Based Network, PBN）上进行多媒体通信的标准，H.324 是在通用交换电信网（General Switched Telecommunication Network, GSTN，包括 PSTN）上进行多媒体通信的标准。

现代的视频会议系统是建立在 ITU-T 的 H.32X 系列框架协议之上的，目前主要以 H.320 和 H.323 应用最为广泛。

H.320 系统是适合电路交换网络的专用解决方案，采用 E1 或 ISDN 线路，专

网专用，确保系统稳定性、可靠性以及图像的高质量。该方案价格高，属于专用系统，适合会场点数少，会议业务量大的企业。从市场机会来看，由于原本就是建立在非规模性应用基础之上，在当前的公众网络资源上用得相对少一些，而且以系统稳定为第一要义，因此这一技术也通常是政府、军队、金融等行业对于安全保障敏感的行业的首要选择。

H.323 系统是适合分组交换网络的开放式解决方案，采用 IP 线路，具有良好的兼容性和互通性。该方案具有组网更加灵活的特点，便于扩展企业内部会场，适合会场点数多、会议业务量大的企业。根据 H.323 协议的规定，一个多媒体会议应用系统，应该由终端、网守（GateKeeper，亦称网闸）、网关（Gateway）、多点控制器（MC）、多点处理器（MP）和多点控制单元（MCU）等部分组成，看似复杂，但用户端所需要做的工作并不多，这也是 TCP/IP 的优势所在。相对而言，基于 TCP/IP 协议的视讯网络部署比较灵活简单，维护也比较方便，所以公众性网络或者企业级视频应用选择 H.323 协议以及相应的 TCP/IP 网络会多一些。

但是由于 IP 网络本身的 QoS 问题还没得到完全解决，视讯技术应用因此容易出现网络抖动、包序混乱、网络拥塞等现象，系统稳定性没有基于 H.320 的视讯系统好。

H.320 因为在国际上已经广泛部署，并且在国内也进行了大量的投资，所以，在未来的多年内还仍然会得到电路交换网络的支持，市场前景并不黯淡；而 H.323 标准则因为和 TCP/IP 绑定到一起，对于系统要求更为简略，并且在操作、互动、以及市场价格方面都具有优势，所以，其发展会更加迅速。

目前，主流视讯设备基本上都是采用 H 系列标准，中国更是如此，电信运营商推出的视讯业务基本上以 H 系列标准为准则，不同的是，或者采用成熟的 H.320 系列协议，或者采用较新的 H.323 系列协议。

除 ITU-T 的 H.32X 之外，IETF 提出的 SIP 协议也是现代视频会议使用的重要标准。SIP 是 IETF 在 1999 年提出一个应用控制（信令）协议，基于文本的应用层控制协议，独立于底层传输协议 TCP/UDP/SCTP，用于建立、修改和终止 IP 网上的双方或多方多媒体会话。SIP 是面向 Internet 应用的设计，同 HTTP、SMTP 等 Internet 应用协议一样，追求实现简单、开放、灵活。SIP 采用文本协议，可扩展性好，通过在报文头域增加新的消息类型来支持新的业务，给厂商留下了足够的扩展空间。

H.323 和 SIP 两者都对 IP 电话系统信令提出了完整的解决方案，但两者的设计风格各有千秋，H.323 采用的是电话信令模式，包括一系列协议；而 SIP 借鉴互联网协议，采用基于文本的协议。当采用 H.323 协议时，各个不同厂商的多媒体产品和应用可以进行互相操作，用户不必考虑兼容性问题；而 SIP 协议应用较为灵活，可扩展性强，两者各有侧重。

在 H. 323 系统中, 终端主要为媒体通信提供数据, 功能比较简单, 而对呼叫的控制、媒体传输控制等功能的实现则主要由网守来完成。H. 323 系统体现了一种集中式、层次式的控制模式。H. 323 是属于国际电联 (ITU) 的标准, 以 H. 323 为标准构建的多媒体通信网很容易与传统 PSTN 电话网兼容, 从这点上看, H. 323 更适合于构建大网。而 SIP 采用 Client/Server 结构的消息机制, 对呼叫的控制是将控制信息封装到消息的头域中, 通过消息的传递来实现。因此 SIP 系统的终端就比较智能化, 它不只提供数据, 还提供呼叫控制信息, 其他各种服务器则用来进行定位、转发或接收消息。这样, SIP 将网络设备的复杂性推向了网络终端设备, 因此更适于构建智能型的用户终端。SIP 系统体现的是一种分布式的控制模式。目前, 有许多运营商正在利用 SIP 构建试验网, 但若想利用 SIP 构建视频大网, 必须对它进行补充、完善, 这样一来 SIP 也不可避免地变得复杂起来。

此外, H. 323 的集中控制模式便于管理, 像计费管理、带宽管理、呼叫管理等集中在控制下实现起来比较方便, 其局限性是易造成瓶颈。而 SIP 的分布模式则不易造成瓶颈, 但各项管理功能实现起来比较复杂, 用户终端操作复杂。

2. 视音频技术发展

视频会议系统早期常用的图像分辨率主要有 QCIF (176×144)、CIF (352×288)、4CIF (704×576)。

CIF 格式受点阵数量的限制, 在数字及线条等精细显示方面不尽如人意, 属于 VCD 的分辨率量级, 画面的流畅度也只有 25 帧 / 秒左右, 视频流畅度非常一般, 远远不能满足目前高标准的要求。虽然 4CIF 的图像在分辨率上有所提高, 但也属于普通 DVD 的分辨率量级, 没有本质的变化。

2006 年初, 第一款 720P 高清视频会议产品问世, 拉开了高清视频会议的序幕。2008 年首款 1080P 高清视频会议系统发布, 标志着视频会议系统已经进入了高清时代。与传统的标清视频会议系统相比, 高清视频会议系统通过提供更为清晰的画面质量、更好的声音效果, 提供给与会者高效、高质量的视频体验, 使与会者能够更有效地进行会议交流。此外, 高清视频会议系统能够高清晰度地显示高分辨率内容, 这对于某些特定领域, 比如, 医疗、地图、测绘等是至关重要的。随着高清电视、高清摄像机等设备的普及, 高清视频会议系统得到了更为广泛的应用, 更多的企业和机构由于认识到信息化建设对于企业的发展起到的至关重要的作用, 对高清视频会议系统有了更多的需求, 同时市场上也涌现出多家厂商, 积极应用高清视频会议新技术, 不断开发出新产品, 整个高清视频会议系统产业氛围已经形成, 视频会议市场已步入高清时代, 进而打开高清视频会议系统的应用新局面。

视频会议系统的图像分辨率从以往的 CIF、4CIF 提高到高清的 720P、1080P，视频图像的帧率也由流畅度非常一般的 25 帧 / 秒、30 帧 / 秒提高到 50 帧 / 秒、60 帧 / 秒，使得视频会议的图像质量有了质的飞跃。目前，高清视频会议已经成为视频会议发展的趋势，标清视频会议将逐渐淡出。

图 1.1 为高清效果对比图，图 1.2 为图像帧率对比图。

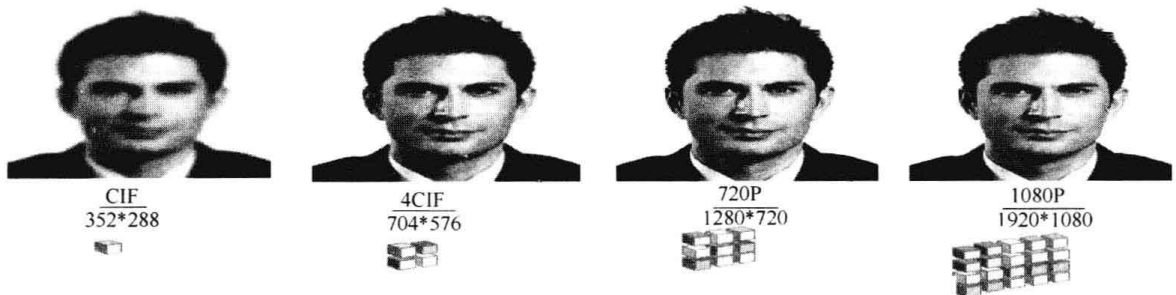


图 1.1 高清效果对比图

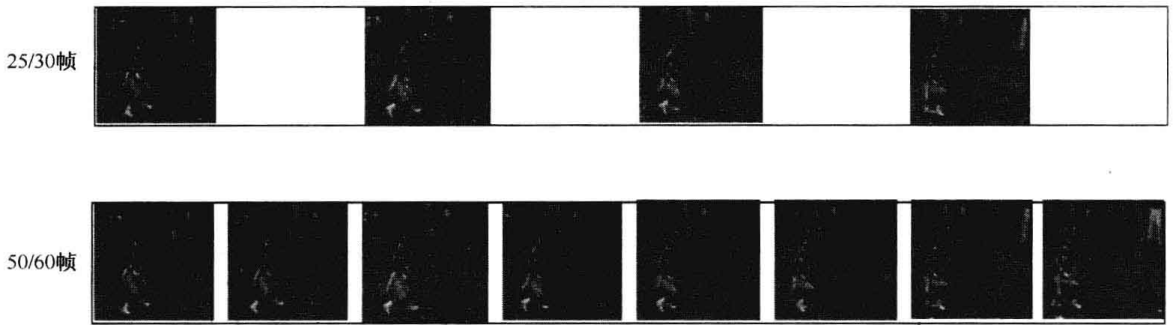


图 1.2 图像帧率对比图

在提高图像分辨率的同时，视频编解码标准从 H.261、H.263 协议发展到主流的 H.264 协议。H.261 视频编解码标准是最原始的视频编解码协议，在 8×8 像素块上使用离散余弦变换（DCT）编码视频帧，支持两种视频图像格式——CIF 和 QCIF，是 H.323/H.320 基本要求的一部分，1995 年之前，会议电视系统图像都采用 H.261 编码协议。

1995 年，ITU-T 针对低比特率视频应用制定了 H.263 标准，当时 H.263 被公认为是以像素为基础的采用第一代编码技术混合编码方案所能达到的最佳结果。尽管采用 H.263 编码技术较 H.261 编码在压缩率和图像质量上都有大幅度的提升，但 H.263 信源编码算法的核心仍然是 H.261 标准中采用的 DPCM/DCT 混和编码算法，原理框图也和 H.261 十分相似。

2001 年 12 月，ITU-T 和 ISO 两个国际标准化组织的有关视频编码的专家联合组成视频联合工作组（Joint Video Team, JVT），负责制定一个新的视频编码标准，以实现视频的高压缩比、高图像质量、良好的网络适应性等目标。随后 JVT 制定出的视频编码标准被 ITU-T 定义为 H.264。H.264 相对以前的编码方法，在图

像内容预测方面提高了编码效率,采用可变块大小运动补偿、1/4 采样精度运动补偿、加权预测等算法,改善了图像质量,增加了纠错功能和各种网络环境传输的适应性。测试结果表明,在中低带宽情况下,H.264 具有比 H.263 更优秀的 PSNR 性能:H.264 的 PSNR 比 H.263 平均要高 3 dB,所以,目前业界主流视频会议系统厂家都推出了基于 H.264 的视讯产品。

H. 264 编码技术使运动图像压缩技术上升到了一个更高的阶段,在较低带宽上提供高质量的图像传输是 H. 264 的应用亮点。

在音频技术方面,视频会议系统中的 AAC-LD 宽带音频协议正逐步取代低品质音频通信协议(G.711、G.722),高保真、低延时宽带音频协议 AAC-LD 成为高品质音频首选

音频协议从 G.711、G.722 发展到宽频语音 AAC-LD。

G.711 标准公布于 1972 年,其语音信号编码是非均匀量化 PCM。语音的采样率为 8 kHz,每个样值采用 8 bit 量化,输出的数据率为 64 kbps。这种窄带编码支持对 300~3 400 赫兹的音频进行压缩,延时长,音质一般,消耗的带宽相对较大,主要用于数字 PBX/ISDN 上的数字式电话。

G.722/G.722.1 标准是用于 1 kHz 采样率的标准化宽带语音编码算法,分别于 1984 年和 1999 年被国际组织定义为标准。G.722/G.722.1 编解码器在 16 kHz 频率上接收 16 位数据(带宽从 50 Hz~7 kHz),并将其分别压缩为 64 kbps 与 48 kbps、32 kbps 与 24 kbps,其总延迟约为 3 ms,能够提供更好的通话质量,但是要求的带宽较宽。G.722/G.722.1 在无线通信系统、VoIP 生产商、个人通信服务、视频会议等领域应用。

AAC (Advanced Audio Coding, 高级音频编码)是由 Fraunhofer 研究院(MP3 格式的创造者)、杜比(DOLBY)试验室和 AT&T(美国电话电报公司)共同研发出的一种音频压缩格式,是 MPEG-2 规范的一部分,并在 1997 年 3 月成为国际标准。随着 MPEG-4 标准在 2000 年成型后, MPEG2 AAC 也被作为核心编码技术,并增加了一些新的编码特性,又叫 MPEG-4 AAC。

MPEG-4 AAC 家族目前共有 9 种编码规格, AAC-LD (Low Delay 低延迟规格)是用于低码率下编码,它支持 48 kHz 采样率,可以 64 kbps 的码率输出接近 CD 高保真音质的音频,并支持多声音通道, AAC-LD 算法延迟为 20 ms。

几种视频会议音频编码协议如表 1.1 所示。

表 1.1 几种视频会议音频编码协议

	采样频率	支持音频带宽	输出码率	最低算法延迟
G.711	8 kHz	300~3 400 Hz	64 kbps	<1 ms
G.722	16 kHz	50 Hz~7 kHz	64 kbps	3 ms

续表

	采样频率	支持音频带宽	输出码率	最低算法延迟
G.722.1	16 kHz	50 Hz~7 kHz	24 kbps、32 kbps	40 ms
G.722.1 .C	32 kHz	50 Hz~14 kHz	24、32、48 kbps	40 ms
AAC-LD	48 kHz	20 Hz—20 kHz	48~64 kbps	20 ms

3. 远程呈现技术

远程呈现技术综合集成了 IP 网络通信、超高清视频编解码、空间 IP 语音、建筑声学、空间照明以及人体工程学等领域的一系列技术创新，从而实现了网络与空间的真实转换，为远在异地的人们营造出一种跨越时空的真实面对面体验。

其产品（如华为公司的智真系统）因其出色的音视频效果、独特的真实体验和感受，并且能够融合统一通信，实现与行业应用的无缝集成，从而具有极大的市场潜力和广阔发展前景。

传统的视频会议体现的是语音和视频的简单交流，而通过远程呈现技术得到真人大小的还原远端的图像，使得会议体验与传统视频会议相比发生了革命性的改变，如图 1.3 所示。



图 1.3 远程呈现的会议效果

人与人的交流不再局限于语音和图像，肢体的动作、精神的状态、眼神的沟通都可以通过远程呈现系统进行捕捉和传达，会议的效果大幅度的提升。

与传统视频会议产品相比，远程呈现产品增加了复杂的多路图像同步、图像拼接处理、环境融合构造等创新技术。通过将 1080P/60 帧极致高清图像进行数字采集，进行像素级的自动拼接处理处理，结合人体工程学的理论，使远方会场中的人物图像以真人 1:1 的比例，按真人大小显示在高清 1080P 显示屏上并辅之以每秒钟 60 帧的高流畅性图像传送，营造出一种与异地对话方共在一个会议室的真实感觉。

4. 良好的网络适用性

为了保护企业的前期投资，视频会议系统应支持标准的 H.320、H.323、SIP 协议，可兼容主流视频会议厂家的设备，提供丰富的第三方接口与企业办公系统进行融合，能够与异构网络的视频系统（如 Microsoft Lync 和支持 TIP 协议的视频会议）融合互通。

为了让企业更方便地与会，同时减少企业的投资成本，视频会议系统应能提供多种接入方式（如 LAN、XDSL、E1），支持有线（如 IP/E1/4E1）和无线（如 3G、WIFI、LTE）等多种网络，融合多种终端达到最强组网效果，让企业随时随地，更方便自如地沟通。

由于 IP 网络存在丢包、延时、抖动等问题，需要具备良好的网络适应性的视频会议系统来保证更好的与会效果，使得企业各分会场能随时随地更加流畅地进行沟通。视频会议系统能够针对 IP 网络传输性能自动调整 QoS 保证技术和采用不同的视音频处理的策略，从而保障图像、音频在各种恶劣网络条件下达到最优效果，在网络丢包达到 5% 时会议效果基本不受影响。

5. 客户化定制

视频会议系统应充分考虑行业化特点，提供客户化的解决方案，同时对于客户提出的对现有系统需要优化或新增功能，提供定制化软件开发，使得视频会议不仅能具备高清的会议功能，还具有丰富的可扩展性功能，满足客户今后不同业务的建设和使用需求。

视频会议系统应能全面融合多种业务，集多媒体通信管理平台，集成视频会议、视频监控、应急指挥调度、即时通信、视频点播、桌面应用、VoIP 电话、办公软件协同等应用于一体，支持多协议的转换和兼容，支持移动网络和 Internet 融合，具有大容量组网、智能网络适应、高保真视音频、软硬结合、多业务融合、平台开放能接入第三方设备等特点。

1.2 视频会议系统的主要功能

1. 多画面会议功能

为了更加真实再现会议的临场效果，视频会议系统应提供强大的多画面功能，会议召开后可灵活选择多画面显示模式，这样各会场都能在一个显示设备上同时显示，极大地增强了会议的临场效果。

多画面会议功能如图 1.4 所示。

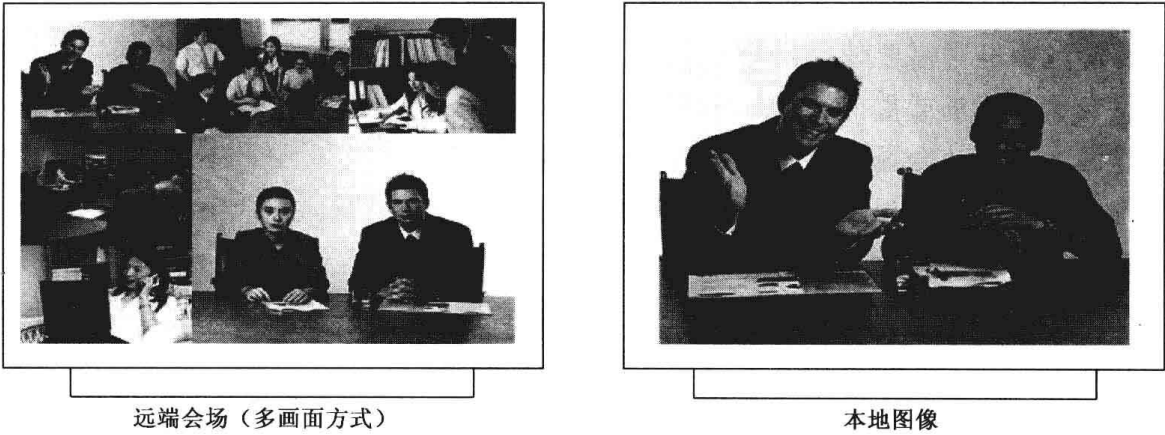


图 1.4 多画面会议功能

2. 双流功能

视频会议中常常需要同时传送演讲人的演讲稿和演讲人的图像或者同时传送两路活动图像。在传送双流图像时，只需要演讲人的 PC 通过 VGA 线缆（适用于标清系统）或者 DVI/HDMI（适用于高清系统）线缆接到会场的终端上，通过标清终端的 VGA 输入接口或者高清终端 DVI/DHMI 输入接口可方便地将演讲稿信息发送到远端会场，远端会场通过标清终端的 VGA 输出接口或者高清 DVI/HDMI 输出接口将演讲稿内容通过投影设备显示出来，而终端的主视频输出接口显示的是远端演讲人的图像。

双流功能如图 1.5 所示。

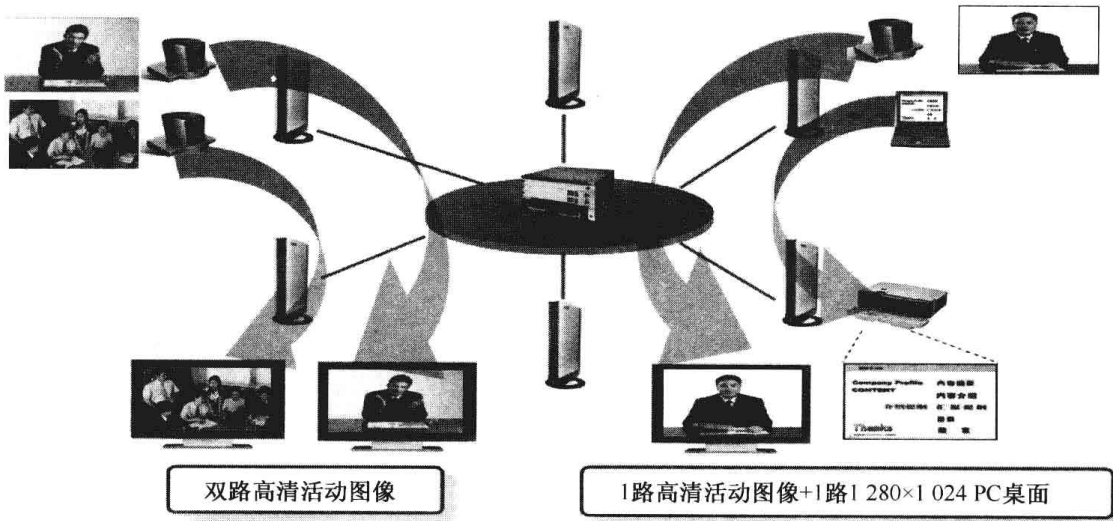


图 1.5 双流功能

3. 字幕与横幅功能

会议中经常需要使用字幕功能，如用于重要提示、会议通知、欢迎词等，如主席会场可现场或提前编辑好字幕，会议中可通过滚动方式或其他方式在会场中实时发送给其他会场。用户可以方便地向远端会场发送上、中、下三个区域的横幅组合内容，并且提供区域的自动调整、自适应实景编辑、图片插入、背景透明度 / 颜色 / 字体可调、多种的字幕滚动方式、中文和英文字幕、预览等强大功能。应用强大的字幕短信功能，增强了与会各会场之间的沟通 and 交流，更贴近用户的使用。

字幕与横幅功能如图 1.6 所示。



图 1.6 字幕与横幅功能

4. 会议召集功能

对于大型正式会议，一般由管理员来调度保证，通过大型管理软件，实现对会议的完全调度、管理、监控和维护，这也是目前业界会议调度的主要方式。此外还应具备其他会议召集功能。

(1) 主叫呼集

会场操控人员仅使用遥控器或 PAD 便可直接呼叫对方的号码（或从电话本选择呼叫），无须其他人协助，可以在任意一个终端会场，自主召集会议，不需要系统管理员协助。系统侧可实现无人值守，实现一键呼集，像打电话一样方便。

(2) 匿名会议（电话会议模式）

匿名会议与电话会议一样，系统预先定义匿名会议后，各会场通过呼叫接入号码、会议 ID 和密码入会。匿名会议对于预先不确定会场的会议非常适合，对于

用户来说，操作方便简单，同时紧贴用户的电话会议使用习惯。

(3) 网络预约

用户可以登录 Web 自行预约会议，到预定时间，系统便自动呼叫终端入会。

(4) 特服号入会

如果采用上述方式召集会议后，某个终端会场未能按时入会，可通过拨打设置的特服号及时入会。

(5) Outlook 预约会议

支持通过 Outlook 预约视频会议，并且预约的会议会自动添加到收件人的 Outlook 日程里，会议开始前系统会自动提醒。

(6) 视音频 IVR 导航创建和加入会议

通过拨打统一接入号或者会议按照视频或者音频提示信息逐步操作，可以创建或者加入会议，或者激活已预定义好会议。

1.3 视频会议系统设计方案

1.3.1 视频会议系统的设计原则

1. 先进性原则

- 系统必须严格遵循国际标准、国家标准和国内通信行业的规范要求；
- 需符合视频技术以及通信行业的发展趋势，并确保采用当前成熟的产品技术；
- 所有的系统采用最先进的技术，确保今后相当长的时间内技术上不会落伍。

2. 开放性原则

- 必须完全符合 H.323 和 H.320、SIP 标准框架协议；
- 必须采用业界标准的视音频编解码协议；
- 必须采用开放式标准设计，兼容标准的视讯系统和设备，确保可与其他厂家标准的产品有效互通；

- 满足今后的发展，留有充分的扩展余地；
- 尽量选择国内主流通信厂家的设备，确保产品得到持续的技术支持和可靠的服务；
- 尽量避免选择小厂家的产品以及采用非标准协议的设备。

3. 可靠性原则

- 确保系统具有高度的安全性；
- 对工作环境要求较低，环境适应能力要强；
- 系统设备安装使用简单，无须专业人员维护；
- 系统需要满足 7×24 小时无人职守方式稳定的工作。

4. 全业务兼容原则

- 要求系统不仅能够提供视讯会议功能，还需要支持丰富的附加业务，满足今后不同业务的建设和使用需求。

5. 经济性原则

- 综合考虑视频会议系统的性能和价格，最经济最有效地进行建设，性能价格比在同类系统和条件下达到最优。

以下以华为公司的视频会议系统产品为例，介绍适合多种典型应用场景下的视频会议系统方案。

1.3.2 华为公司视频会议系统解决方案

1. 总体架构

整个系统采用 IP 传输网络，基于标准的 H. 323 架构，是一个开放的系统，视频编码采用业界主流的 H. 264 编码，同时兼容 H. 261、H. 263 编码，可以为用户提供 1080P 高清视频图像。配合最新的 H. 264 HP 编码技术，大大提高了图像压缩比，在同样带宽下，可向用户提供更逼真、更清晰、更流畅的画面。根据用户的实际配置需求，可以提供 50/60 fps 技术，在同等带宽下，将屏幕刷新频率增加一倍，增强图像的流畅度。可以提供动高清双流，实现静态和动态的高清晰图像。

音频编解码协议采用 AAC-LD/LC，该协议采用 48 kHz 采样率，码率最高达 96 kpbs，并且提供多声道语音支持，给实时通信带来 CD 级的音质效果。

系统支持 H.239 高清双视频流和 SXGA 输入 / 输出特性，主流支持 1080@60fps，

辅流支持 1080P@30fps, 用户直接通过带内的方式就可以实现全景式会议, PowerPoint 演讲稿与 Word 演讲稿可以同步显示。

整个系统具有高可靠性, 支持诸如 IP 网口备份、N+1 备份、电源备份、芯片备份, 保证视频会议系统全天 24 小时, 全年 365 天无故障运行。

2. 视频会议系统典型解决方案

在总部中心机房部署华为高清 MCU、SIP 服务器、会议管理平台 and GK, 全网由中心点会议管理平台统一管理, 全网资源统一调度, 在分支机构控制中心机房配置华为高清 MCU, 总部和分支机构的 MCU 实现级联。

华为 MCU 支持数字级联, 且最大支持 5 级级联。数字级联采用无缝对接, 主会场的信号经过两级 MCU 转发给下级分会场, 中间无须编解码处理, 视音频效果未受损伤, 除转发时网络延时外, 跟使用同一台 MCU 毫无差别。

华为 MCU 的数字级联技术可以使会场操控更方便简单。两 MCU 间有严格的上下级关系, 上级 MCU 进行广播操作后, 下级 MCU 会自动进行广播, 主会场发送辅流后, 辅流也会自动发送给下级的所有会场, 统一维护调度无须额外操作, 级联时间不超过 5 s。

华为 MCU 支持多通道级联, 且设有快捷的操作方式, 充分降低管理人员的运维压力。系统根据 MCU 的级联多通道的线路类型和数量, 自动调度出多通道级联会议, 并自动生成多通道参数信息, 使用户方便自由选看任何会场, 带来更好的会场体验效果。

通过采用华为 MCU 级联召集会议的方式, 可以解决分层组网、提高会议容量等问题, 如图 1.7 所示。此外, 华为高清 MCU 支持 1080P、720P、4CIF、CIF 视频会议终端接入, 支持 ITU-T H.323、H320 和 IETF SIP 协议, 支持多点智真, 支持智真会场、高清、标清的混合会议。

华为会议管理平台 SMC2.0 负责会议的召集、管理和结束等会议管理控制功能, 通过会议管理平台对这些码流进行分发、交换、会议控制等。

华为 GK 服务器负责会议的注册, 同时具备 H.460 功能的专业安全服务器。

SIP 服务器负责建立网络中所有的 SIP 电话通话。整体网络结构如图 1.8 所示。

对于各地分会场, 根据实际需求可分别配置高端会议室、复合型多功能会议室、桌面会议室和个人终端。

高端会议室可部署华为沉浸式智真, 满足最大 18 人与会, 提供单排或双排智真选择, 使用单个或 3 个 65 寸专业显示屏, 具备 1080P@60fps 的极致高清视频、三声道音频、真人大小、听声辨位等特点, 达到虚拟的面对面高临场感沟通, 帮助企业 CXO 们或中高层管理人员的及时沟通, 高效协作, 快速决策。

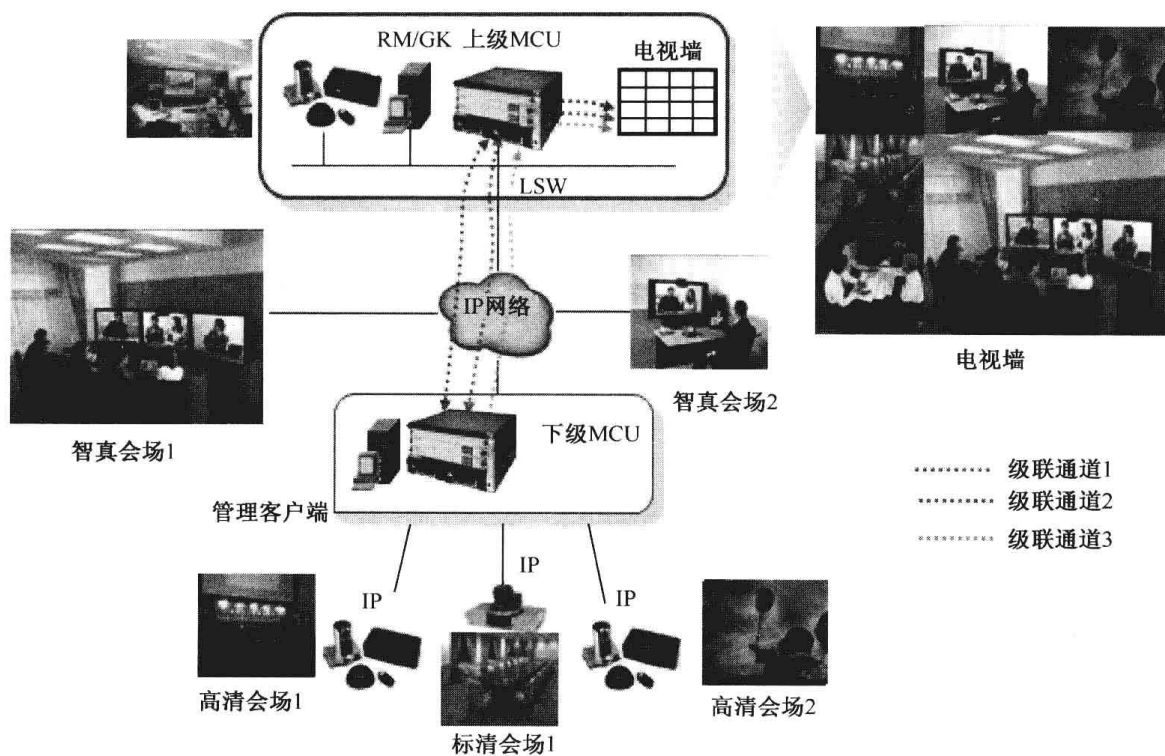


图 1.7 MCU 多通道级联图

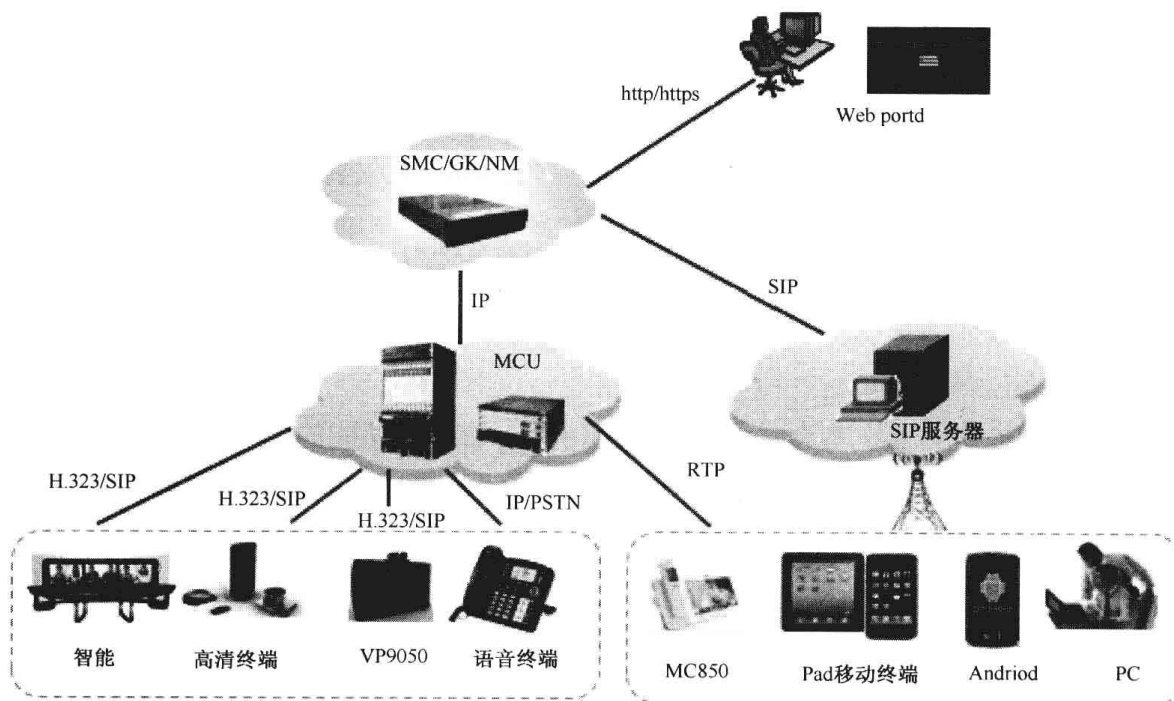


图 1.8 华为公司典型视频系统建设组网图

复合型多功能会议室部署华为全系列分体式高清终端，包括华为 ViewPoint

9000 系列高清会议终端、高清摄像机、高清阵列式麦克风，同时企业需要自行配置大屏幕液晶电视、投影仪等设备。华为高清会议系统具备业界最强的 H.264 1080P@60fps 帧视频编解码技术，宽频语音（AAC-LD）和高清晰数据传送，最大可支持 8 Mbps 会议带宽，提供 5 倍 DVD 画质高清晰图像和高保真语音，让企业与会者全方位体验极度高清的视频会议。

桌面会议室部署华为宽带多媒体会议终端。华为 MC850 系列多媒体电话集宽带电话、数码相框、视频娱乐于一身，创新企业数字生活，重新定义企业桌面通信平台。MC850 系列宽带多媒体电话具备 7 英寸 TFT 液晶触摸屏、数字无绳手柄、内置摄像头，集成 TV（CVBS 复合视频）输出、Ethernet 方式接入内置 Web 浏览器、SD/MMC 卡扩展，让企业员工随时随地、方便快捷地沟通，提高了企业工作效率。

个人终端设备配置安装华为高清软终端，满足个人临时会议的需求。华为高清软终端可以满足桌面型（Windows, Mac）及移动型（Android, IOS）需求，无缝接入视频会议系统和远程呈现系统，视频、音频、数据融合会议，随时随地加入会议，方便易用。

3. 视频会议的多种接入方式

为提高用户随时随地接入视频会议系统的方便性，华为公司的视频会议产品支持多种接入方式。

（1）E1 专线接入

华为高清视频会议系统提供 IP、4E1 等多种线路传输方案，方案除必配的 IP 接口外，4E1 接口可通过接口板的方式接入，即插即用，可以最大程度上满足用户的不同接入需求，特别是已有 E1 专线的用户，更能节约成本，保护原有投资，如图 1.9 所示。

（2）3G/LTE 数据卡无线接入

通过把华为 3G/LTE USB 数据卡插入华为高清终端 USB 接口，华为高清终端可自动检测到数据卡，并建立 USB 传输通道，高清终端通过拨号获取到 IP，使用该 IP 即完全可像有线环境一样进行相关网络操作，如注册 GK。

本端图像和声音经过编码形成码流，码流通过 USB 接口传输给 3G/LTE USB 数据卡，3G/LTE USB 数据卡通过无线信号（3G/LTE）发送数据，远端接收到码流解码出图像和声音，从而达到实时沟通的效果。

3G/LTE USB 数据卡无线接入解决方案示意如图 1.10 所示。

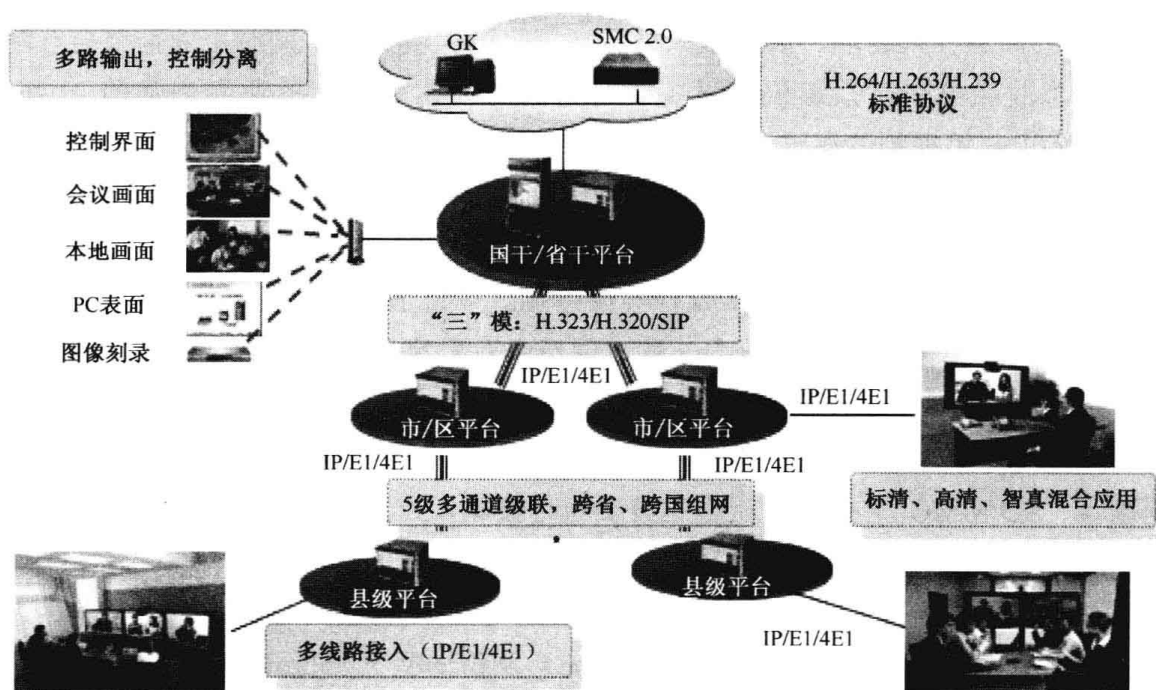


图 1.9 E1 接入方式

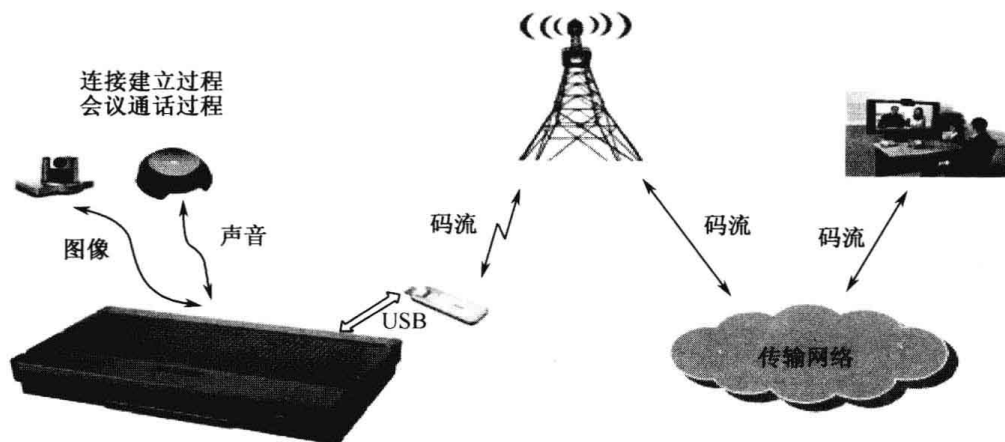


图 1.10 3G/LTE USB 数据卡无线接入解决方案示意图

(3) 卫星接入

卫星网络充分解决地域通信的限制, 让语音、视频、数据能在较偏远山区以及有线信号无法到达的区域进行传输。华为视频会议系统支持 H.323、H.320 两种框架协议, 适应卫星传输的特点, 能在低带宽下保证较好的图像效果, 同时卫星地面接收到信号以后, 转换为 ISDN、E1 等线路就能直接接入终端和 MCU, 非常方便地利用线路特点, 实现方式简单。该方案为用户分支点的接入提供了无限的可能, 完善了所有线路的接入方式。图 1.11 为卫星接入解决方案示意图。

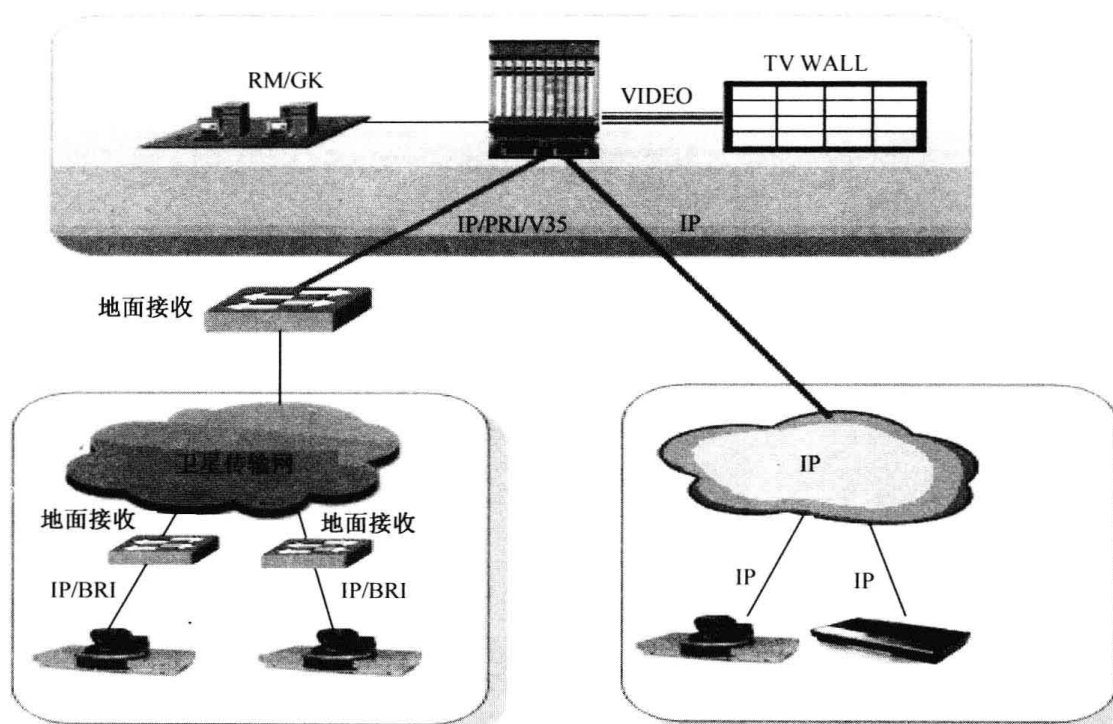


图 1.11 卫星接入解决方案示意图

(4) 支持 PSTN 以及移动电话语音接入，满足在外人员接入视频会议需求

华为公司的 MCU 完全支持 PSTN 电话和移动电话语音接入，对于平台的会议管理不需要任何更改，通过 IP 语音网关方式，接入企业的 PSTN 电话和移动电话，可实现企业内的任意电话接入，实现了视频、语音的最大兼容。

第2章 光度与色度

光度和色度是视频的基础，在开始介绍视频会议技术之前，需要对光的基本知识有一个系统的了解。

本章知识要点如下：

- 光度学知识；
- 人眼的视觉特性；
- 色度学知识。

2.1 光度学知识

2.1.1 可见光

由光学理论知道，光是一种以电磁波形式存在的物质，它的波长范围大约在380~780 nm（1 nm=10⁻⁹ m）之间。由于这段电磁波的辐射能为人眼看得见，所以称为可见光，习惯上简称为光。人眼对不同波长的光引起不同的颜色感觉，例如，波长为400 nm左右和波长为700 nm左右的光，给人以紫和红的感觉。在可见光的范围内，按波长的依次递减，相应颜色排列为红、橙、黄、绿、青、蓝、紫七种颜色，把这些色光混合在一起就得到白光。

图2.1给出了电磁辐射波谱和可见光的波长范围。

2.1.2 常用的光学物理量

1. 色温

色温用于表示光源的光谱性能，其定义为在可见光谱内，当光源的光谱等同于某温度的绝对黑体辐射的光谱时，则该绝对黑体的温度称为该光源的色温，单位为绝对温度开氏度（K）。

我们知道，温度在绝对零度（-273℃）以上的物体都会有连续的电磁辐射。但是不同的物体的辐射能量是不同的。为了衡量物体的电磁辐射能量的大小，人

们设定了一个标准——绝对黑体。绝对黑体是指在任何温度下，对于各种波长的电磁辐射的吸收系数恒等于1 的物体。自然界并不存在绝对黑体。绝对黑体是一个理想化的参考模型，实验模型是一个中空的、内部涂黑的球体，其上面开一个小孔，这个小孔就相当于绝对黑体。随着温度的增加，绝对黑体辐射能量增大，其功率波谱的峰值向短波方向移动。

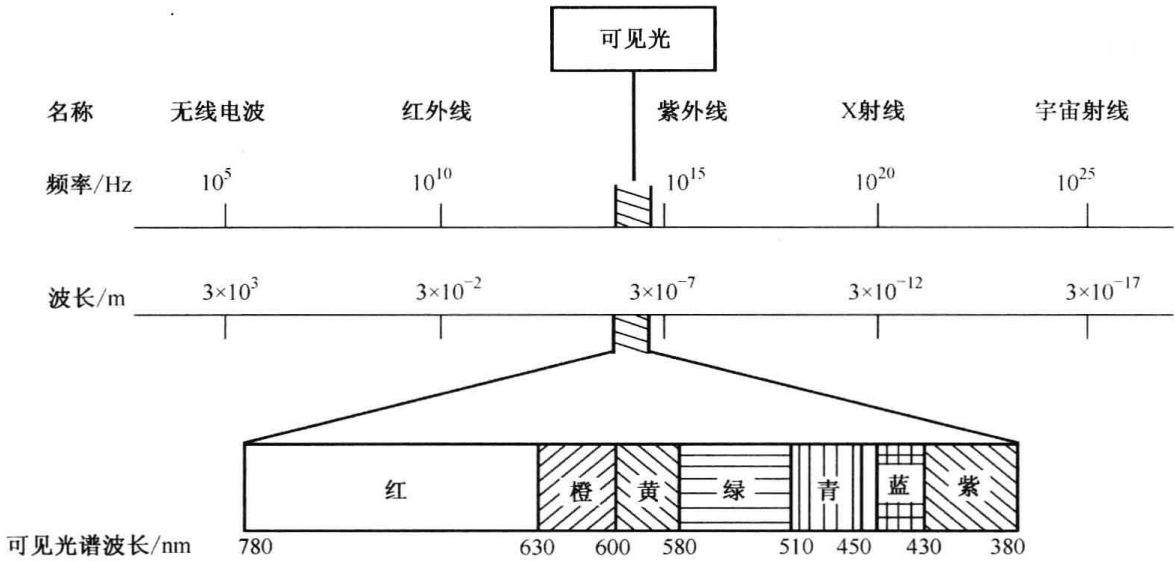


图 2.1 电磁辐射波谱和可见光的波长范围

需要注意的是，色温与光源的实际物理温度无关，它是表征光源波谱特性的参量。如温度为 2 800 K 的钨丝灯发出的白光，色温为 2 854 K。

为了对色温有更加实际的感性认识，表 2.1 给出了一些生活中常见光源的实际色温值。

表 2.1 生活中常见光源的色温值

光源	色温 / K	光源	色温 / K
油灯	1 900~2 000	白炽灯（10W）	2 400
蜡烛	1 900~2 000	白炽灯（40W）	2 700
月亮	4 100	白炽灯（100W）	2 750
日出后及日落前太阳光	2 200~3 000	白炽灯（500W）	2 900
中午太阳光	4 800~5 800	日光色荧光灯	6 500
下午太阳光	4 000	冷白色荧光灯	4 300
阴天自然光	6 400~6 900	暖白色荧光灯	2 900
晴朗的蓝天	8 500~22 000	普通高压钠灯	2 000

国际照明委员会（简称 CIE）也规定了一些标准光源，如表 2.2 所示。

表 2.2 标准光源色温值

光源	色温 / K	描述
A 光源	2 854	相当于 2 800 K 钨丝灯发出的光，其光源能量主要集中在波长较长的区域，因此，光总带着橙红色
B 光源	4 874	相当于中午直射的太阳光
C 光源	6 774	相当于白天的自然光，被选为 NTSC 制彩色电视系统的标准白光源
D ₆₅ 光源	6 504	相当于白天平均照明光，被选为 PAL 制彩色电视系统的标准白光源
E 光源	5 500	一种理想的等能量的白光源，实际并不存在

2. 光通量

光通量指人眼所能感觉到的辐射能量，它等于单位时间内某一波段的辐射能量和该波段的相对视敏度的乘积。由于人眼对不同波长光的相对视敏度不同（2.2 节将会介绍），所以不同波长光的辐射功率相等时，其光通量并不相等。例如，当波长为 555 nm 的绿光与波长为 650 nm 的红光辐射功率相等时，由于绿光的相对视敏率是红光的 10 倍，所以绿光的光通量为红光的 10 倍。

在各色光中，黄、绿色光能激起最大的明亮感觉。如果用绿色光作为水准，令它的光通量等于辐射能量，则对其他色光来说，激起明亮感觉的本领比绿色光小，其光通量也小于辐射能通量。光通量的单位是流明，是英文 lumen 的音译，简写为 lm。

3. 发光强度

发光强度简称光强，用符号 I 表示。国际单位是 candela（坎德拉），简写 cd。1 cd 是指光源在指定方向的单位立体角内发出的光通量。

如果光源辐射是均匀时，则光强为 $I=F/\Omega$ 。 Ω 为立体角，单位为球面度（sr）； F 为光通量，单位是流明。

4. 光亮度

光亮度是表示发光面明亮程度的，指发光表面在指定方向的发光强度与垂直且指定方向的发光面的面积之比，单位是坎德拉 / 平方米（cd/m²）或尼特（nt）。

5. 光照度

光亮度是用于测量主动发光物体的物理量，光照度是用于测量被动照射物体的物理量，比如环境光的大小。光照度的单位是勒克斯，也可写为 lx。1 勒克斯的光照度等于 1 流明的光通量均匀分布在 1 平方米面积上被照射的物体上。以下是

各种环境光照度参考值：黑夜为 $0.001\sim 0.02\text{ lx}$ ，月夜为 $0.02\sim 0.3\text{ lx}$ ，阴天室内为 $5\sim 50\text{ lx}$ ，阴天室外为 $50\sim 500\text{ lx}$ ，晴天室内为 $100\sim 1000\text{ lx}$ ，夏季中午太阳光下的照度为 10^9 lx 。

2.2 人眼的视觉特性

2.2.1 人眼对光亮度的感觉

现代神经生理学实验证实，人眼的视网膜上存在大量的光敏细胞，按其形状可分为两类，即杆状细胞和锥状细胞。杆状细胞对入射光的强度很敏感，它能分辨亮度的差别，而对颜色的分辨本领差；锥状细胞虽然对亮度的灵敏度不高，但却能分辨颜色。白天亮度较强时，主要靠锥状细胞产生视觉作用，所以白天人所看到的是万紫千红的彩色世界。夜间亮度较弱时，主要靠杆状细胞产生视觉作用，因此夜间看到的物体无颜色之分，只有灰蒙蒙的影像。归纳而言，杆状细胞决定了人眼的暗视觉特性，锥状细胞决定了人眼的明视觉特性。

大量实验表明人眼对不同波长光亮度感觉的灵敏度是不同的，这称为视敏度。国际照明委员会（CIE）正式推荐的明视觉光谱光效率曲线 $V(\lambda)$ 与暗视觉的光谱光效率曲线 $V'(\lambda)$ ，如图 2.2 所示， $V(\lambda)$ 与 $V'(\lambda)$ 函数表如表 2.3 所示。

表 2.3 明视觉与暗视觉的光谱光效率函数（最大值为 1）

波长 / nm	明视觉 $V(\lambda)$	暗视觉 $V'(\lambda)$	波长 / nm	明视觉 $V(\lambda)$	暗视觉 $V'(\lambda)$
380	0.000 04	0.000 589	530	0.862	0.811
390	0.000 12	0.002 209	540	0.954	0.650
400	0.000 4	0.009 29	550	0.995	0.481
410	0.001 2	0.034 84	560	0.995	0.328 8
420	0.004 0	0.096 6	570	0.952	0.207 6
430	0.011 6	0.199 8	580	0.870	0.121 2
440	0.023	0.328 1	590	0.757	0.065 5
450	0.038	0.455	600	0.631	0.033 15
460	0.060	0.567	610	0.503	0.015 93
470	0.091	0.676	620	0.381	0.007 37
480	0.139	0.793	630	0.265	0.003 335
490	0.208	0.904	640	0.175	0.001 497
500	0.323	0.982	650	0.107	0.000 677
510	0.503	0.997	660	0.061	0.000 312 9
520	0.710	0.935	670	0.032	0.000 148 0

续表

波长 / nm	明视觉 $V(\lambda)$	暗视觉 $V'(\lambda)$	波长 / nm	明视觉 $V(\lambda)$	暗视觉 $V'(\lambda)$
680	0.017	0.000 071 5	740	0.000 25	0.000 001 379
690	0.008 2	0.000 035 33	750	0.000 12	0.000 000 760
700	0.004 1	0.000 017 80	760	0.000 06	0.000 000 425
710	0.002 1	0.000 009 14	770	0.000 03	0.000 000 241 3
720	0.001 05	0.000 004 78	780	0.000 015	0.000 000 139 0
730	0.000 52	0.000 002 546			

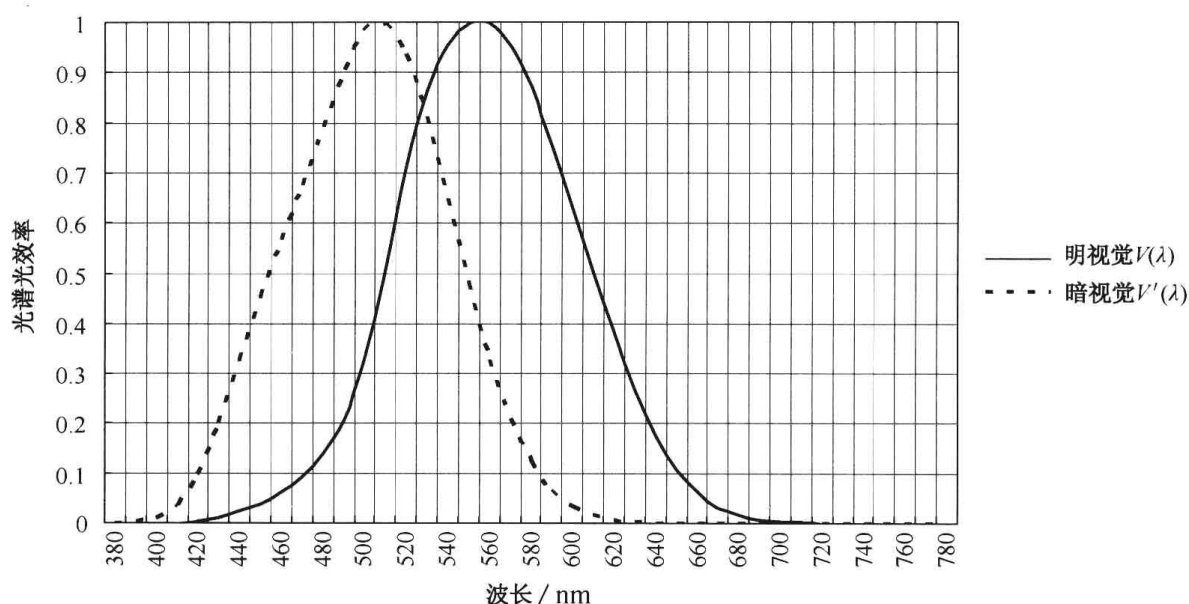


图 2.2 明视觉与暗视觉的光谱光效率

从图 2.2 可以看出,光源的某些波长不在人眼的光谱光效率特性函数 $V(\lambda)$ 之内,人眼是无法感觉的,例如,红外光和紫外线。在人眼可以感觉的部分,人眼对各个波长的光感也是不同的。

2.2.2 人眼对光色度的感觉

人眼的视觉系统中存在三种光谱特性不同的光敏组织 $V_R(\lambda)$ 、 $V_G(\lambda)$ 、 $V_B(\lambda)$,分别对应红(R)、绿(G)、蓝(B)三种锥状细胞,其最大值分别在光谱的红、绿、蓝区域,如图 2.3 所示。

当光作用到视网膜上时,虽然能同时引起三种光敏细胞的兴奋,但由于光的波长和光的强度不同,人们感觉到的颜色不同。例如,红、绿、蓝三种光敏细胞的红光敏细胞感觉最为强烈,则产生红色的感觉;如果同时引起红、绿光敏细胞的兴奋,就会产生黄色的感觉;如果某种光能引起红、绿、蓝三种光敏细胞同样强烈的兴奋,就会产生白色的感觉。

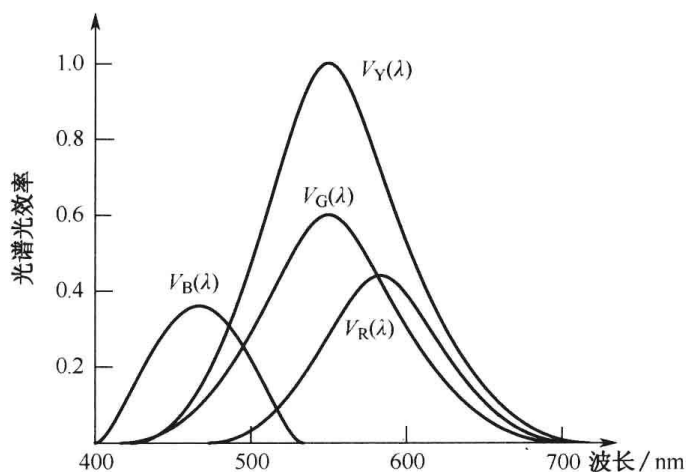


图 2.3 三种锥状细胞的光谱光效率曲线

当产生某一颜色的感觉时，三种光敏细胞或多或少均会产生兴奋，所以每种颜色都会有白色光成分，即有亮度的感觉。三种光敏细胞的兴奋比例程度不同，就会产生各种不同的颜色的感觉。某种复合光的光亮度等于红、绿、蓝光亮度之和。

一般来说，如果两种光谱成分不同的光，作用到视网膜产生的光感与单色光作用到视网膜产生的光感相同，则人们就认为合成光与单色光的彩色感觉相同，是同一种颜色，这就是同色异谱光的概念。

人眼根据红、绿、蓝三基色在视网膜上的光刺激的混合比例来确定颜色感的方法称为混色法。混色法有相加混色和相减混色之分。

相加混色法，即以彩色光的互相叠加来实现混色产生另一种新的彩色光。相加混色法的混色规律如图 2.4 表示，由图可见，以等量的红、绿、蓝三基色光进行相加混色效果如下：

红色+绿色=黄色；

绿色+蓝色=青色；

蓝色+红色=紫色；

红色+绿色+蓝色=白色。

相加混色又分为时间混色法、空间混色法和生理混色法。时间混色法是三种基色轮流投射到某一表面，轮换时间足够快，利用人眼视觉惰性，得到相加混色的效果。空间混色法是利用人眼空间分辨力有限的特性，将近距离三基色发光点看成相加混色。生理混色法是利用两眼同时观看两种不同颜色的同一幅图像，可以得到混色的效果。

彩色电视系统中使用的是相加混色法，图 2.4 为红、绿、蓝三基色混色的原理图。

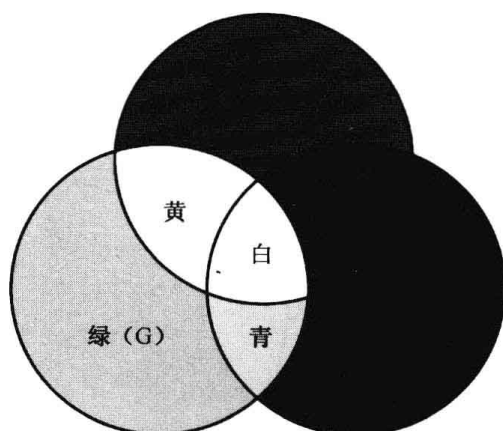


图 2.4 红、绿、蓝三基色混色的原理图

相减混色法是利用颜料、染料吸色性质实现的。印刷、美术等行业使用的是相减混色法。

人眼对不同波长的光有不同的色调感觉，人眼能分辨出色调差别的最小波长变化称为色调分辨阈，其数值随波长变化，如图 2.5 所示，色调分辨力与色调分辨阈成反比。

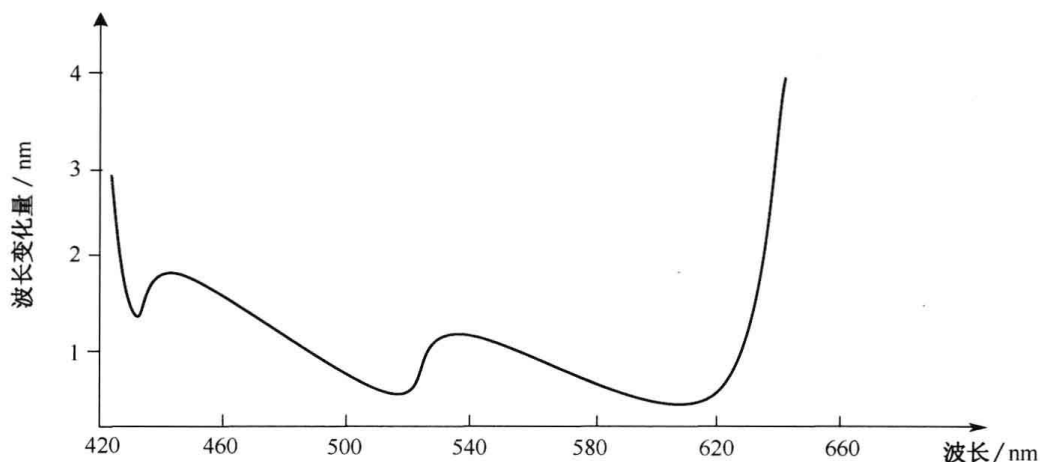


图 2.5 色调分辨力与波长的关系

由图 2.5 可知，人眼对 480~640 nm 区间的光的色调分辨力较高，其中对 500 nm（青绿）和 600 nm（橙黄）两个波长来说，只要波长变化 1 nm，便可以分辨出颜色色调的变化；而有的波长区间，例如，655 nm 的红色到可见光谱长波的长端，以及 430 nm 的紫色到可见光谱的短端，人眼几乎看不到色调的差别。当色饱和度减小时，人眼的色调分辨力将下降；当亮度太大或太小时，人眼的色调分辨力也会下降。

人眼对不同颜色的色饱和度的分辨力能力较色调分辨能力小，并随波长的变化而不同。实验表明，在 580 nm 的黄色区，可以分辨出 4 级饱和度；而在红、蓝区域，可分辨出 25 个色饱和度，如图 2.6 所示。

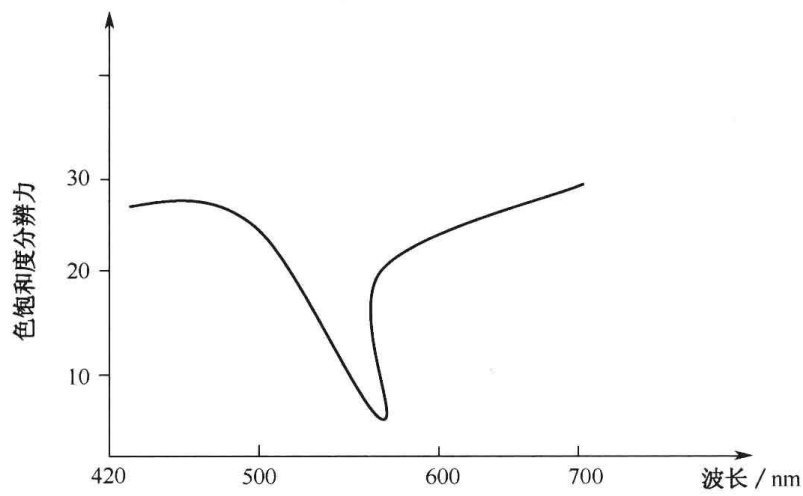


图 2.6 饱和度分辨力与波长的关系

正因为人眼对色调失真敏感，对色饱和度失真不敏感。因此，在模拟电视系统中，必须设置色调调整功能，以消除因传输或解调误差造成的色调失真。而在数字电视系统中，其色调失真很小。

2.2.3 人眼的分辨力

当与人眼相隔一定距离的两个黑点靠近到一定程度时，人眼就分辨不出有两个黑点存在，而只感觉到是连在一起的一个点。这种现象表明人眼分辨景物细节的能力是有一定极限的。我们可以用视觉锐度来定义人眼的分辨力。视觉锐度即人眼对被观察物体刚能分辨出它上面最紧邻两黑点或两白点的视角，如图 2.7 所示。

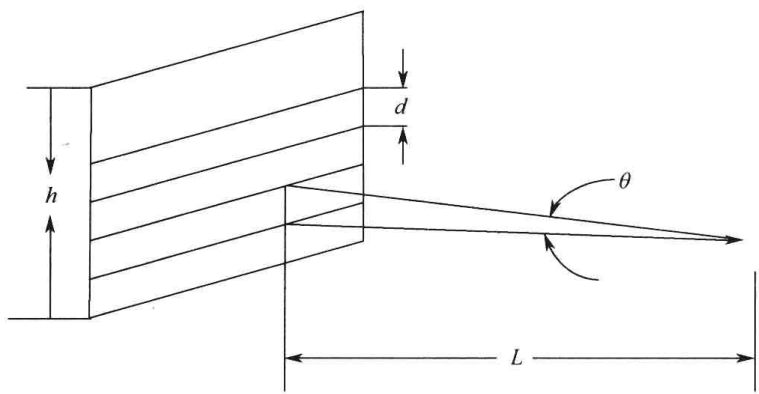


图 2.7 人眼视觉锐度示意图

在图 2.2.6 中， L 表示人眼与图象之间的距离， d 表示能分辨的最紧邻两黑点之间的距离。 θ 为视觉锐度，表示观测点（眼睛所在点）与人眼能分辨的相距最近的两个点所形成的夹角。视觉锐度 θ 的倒数称为视敏度。通过简单的几何运算，可以得到下列公式：

$$\frac{d}{2\pi L} = \frac{\theta}{360 \times 60}$$
$$\theta \approx 3438 \frac{d}{L} \quad (2.1)$$

可见，视觉锐度 θ 越小，分辨图像细节的能力越强。正常人的视觉锐度一般在 $1' \sim 1.5'$ 之间。视觉锐度一般与下列因素有关：

(1) 图像平均亮度越亮，人眼的视觉锐度 θ 越小，分辨图像细节的能力越强

这主要是因为在较亮的环境下，人眼的视觉特性主要取决于明视觉特性，明视觉特性主要是锥状细胞在起作用，而人眼的锥状细胞具有较高的分布密度，因此比较容易分辨图像细节。而在较暗的环境下，人眼的视觉特性主要取决于暗视觉特性，暗视觉特性主要是杆状细胞在起作用，而人眼的杆状细胞具有较疏的分布密度，因此分辨图像细节的能力较差。

(2) 图像的对比越强，对人眼的刺激越大，分辨图像细节的能力越强

如果黑白相间的图像混为一片灰色，失去对比，就无法分清图像的原始面貌。

(3) 人眼对静止图像的分辨力高于运动图像的分辨力

这是由于人眼存在视觉惰性，光刺激在视网膜上会存储一段时间，下一个光刺激可能与上一个光刺激向重合，造成图像分辨力下降。

(4) 人眼对黑白图像的分辨力高于彩色图像

统计结果表明，人眼的彩色视觉锐度比黑白的视觉锐度大 $3 \sim 5$ 倍。如果把人眼的黑白颜色组合细节的分辨力定为100%，则其他颜色的分辨力如表2.4所示。

表 2.4 人眼对彩色细节的分辨力

细节色别	黑白	黑绿	黑红	黑蓝	绿红	红蓝	绿蓝
分辨力	100%	94%	90%	26%	40%	23%	19%

根据这一理论，人们在对电视图像进行处理时，可以对图像亮度信号进行比较细致的处理，如保证其模拟带宽或保证其信号采样率，而对色度信号适当降低带宽或信号采样率。当亮度信号和色度信号重合时，人眼的感觉仍然是图像清晰、颜色鲜艳的彩色画面，这就是所谓的大面积图像着色理论。

2.2.4 人眼视觉惰性

当一定强度的光突然作用于视网膜时，人眼并不能立即产生稳定的亮度感觉，而须经过一个指数规律的变化过程才能达到稳定。另外，当作用于人眼的光线突然消失后，亮度感觉并非立即消失，而是近似按指数规律下降而逐渐消失的，如图 2.8 所示。

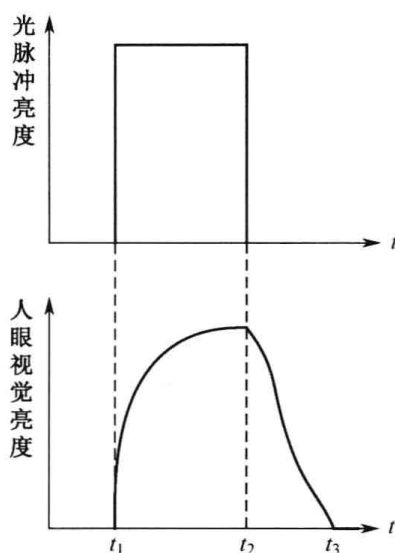


图 2.8 人眼视觉惰性

当光线消失后的视觉残留现象称为视觉暂留或视觉残留。人眼视觉暂留时间为 $t_3 \sim t_2$ 。在中等亮度刺激下，正常视力的人视觉暂留时间为 0.1 s。人眼亮度感觉变化滞后于实际亮度变化，以及视觉暂留特性，总称为视觉惰性。视觉惰性是现代电影和电视技术的理论基础之一。

当人眼收到频率较低的周期性的光脉冲刺激时，就会有明暗交替的亮度感觉，这种现象称之为闪烁。如果光脉冲的频率增加到一定程度后，闪烁就会消失，并给人以稳定、连续、流畅的感觉。人们把人眼恰好感觉不到闪烁的频率称之为临界闪烁频率 (Critical Fusion Frequency, CFF)。临界闪烁频率 f_c 与下列因素有关：

(1) f_c 与亮度的关系

$$f_c = a \lg L_m + b \quad (2.2)$$

式中， a 、 b 为常数，一般 $a=9.6$ ， $b=26.6$ ， L_m 为屏幕亮度，单位为 cd/m^2 。当 $L_m=100 \text{ cd/m}^2$ 时， $f_c=45.8 \text{ Hz}$ 。当屏幕亮度上升时，临界闪烁频率 f_c 随之上升。由临界闪烁频率可以确定广播电视系统的图像场扫描频率，即场频或图像刷新频率。NTSC 制为 60 Hz，PAL 制为 50 Hz。

(2) f_c 与图像显示面积的关系为

$$f_c = c \lg A + d \quad (2.3)$$

式中, A 为图像的显示面积, c 、 d 为常数。上式说明图像的显示面积越大, 人眼的临界闪烁频率 f_c 越高。这样对于大屏幕的图像显示, 必须提高其场频, 可以减小大面积的图像闪烁。

(3) f_c 与图像颜色的关系

白光的亮度最大, 临界闪烁频率 f_c 最大; 其他依次为绿、红、蓝。

2.2.5 人眼的视觉适应性与图像对比度

人眼对不同的环境光具有不同的亮、暗适应性, 对于不同的亮度 B 能观察到的最小亮度变化 $\Delta B_{\min}$ 也是不同的。但在相当大的范围内。可以觉察的最小相对亮度变化 $\Delta B_{\min}/B$ 等于常数 δ 。常数 δ 称为对比灵敏度阈值。对比灵敏度阈值与环境光的大小有关系, 通常在 0.005~0.02 范围内。对于视觉刺激来说, 假定 $\delta=0.01$, 表示如果原亮度为 100 cd/m^2 , 亮度变化 1 cd/m^2 , 人眼就可以察觉出亮度变; 如果原亮度为 1 000 cd/m^2 , 亮度需要变化 10 cd/m^2 , 人眼才可以察觉出亮度变化。因为人眼对亮度有一定的适应性, 高亮度 1 000 cd/m^2 和低亮度 100 cd/m^2 的对比灵敏度阈值 δ 并不变化, 这种适应性是由人眼瞳孔大小调节和视网膜的感光物质变化形成的。

这种特性对实际传送图像信号十分重要, 我们无须传送实际的图像亮度范围, 只需传送、重显一定对比度 C 的图像, 即图像的最大亮度 $B_{\max}$ 与最小亮度 $B_{\min}$ 之 $C=B_{\max}/B_{\min}$ 不变就可以了。

图像的亮度范围由景物的反射系数范围决定。在一定照度下, 最白的白石膏反射系数接近于 1.0, 最暗的黑丝线的反射系数为 0.01, 所以被传图像的对比度不会超过 100。

图像对比度对重显图像质量至关重要, 适当的图像对比度可以使图像层次分明, 因此人们总是想尽办法提高图像的对比度。但也不是无限提高图像的对比度, 因此正常景物的对比度只有 100, 因此过高的对比度并不能改善图像质量。

例如, 被摄图像的亮度范围为 10~1 000 cd/m^2 , 对比度为 100。传输到显像端的亮度范围为 1~100 cd/m^2 , 对比度也为 100。人眼感觉到图像与摄像图像基本相同。电视系统巧妙地利用了人眼的视觉适应性, 只需传送一定对比度的图像, 不用传输实际的亮度范围, 人们仍然可以看到层次分明的图像。

2.3 色度学知识

2.3.1 彩色三要素

物体或图像的亮度、色调、色饱和度被称作物体或图像的彩色三要素，其中色调和色饱和度合称色度。

亮度是指色光的明暗程度，它与色光所含的能量有关。对于彩色光而言，彩色光的亮度正比于它的光通量。对物体而言，物体各点的亮度正比于该点反射（或透射）色光的光通量大小。一般来说，照射光源功率越大，物体反射（或透射）的能力越强，则物体的亮度越强。

色调是指颜色的类别，通常所说的红色，绿色，蓝色等，就是指色调。光源的色调由其光谱分布决定，物体的色调由照射光源的光谱和物体本身反射特性或者透射特性决定。例如，蓝布在日光照射下，只反射蓝光而吸收其他成分。如果分别在红光、黄光或绿光的照射下，它会呈现黑色。红玻璃在日光照射下，只透射红光，所以是红色。

饱和度是指颜色的深浅程度，即颜色的浓度。对于同一色调的彩色光，其饱和度越高，它的颜色就越深；饱和度越低，它的颜色就越浅。某一色调的彩色光中掺入白光，其饱和度下降，掺入的白光越多，其饱和度就越低。

2.3.2 1931 CIE-RGB 计色系统

为了描述、计量、测试颜色的特性，并使国际间不同实验室的数据可以相互比较，需要选择一个公认的色度计量体系。

根据 2.2.2 节的论述，为了准确的对颜色进行计算，CIE 规定红、绿、蓝三基色波长各为：

- 标准红基色的波长选为 700 nm；
- 标准绿基色光选为 546.1 nm；
- 标准蓝基色光为 435.8 nm。

由于这三种基色光可以用物理手段产生出来，因此通常称为物理三基色，它们是色度学计量中最基本的一套三基色。

CIE 规定将能匹配出标准白光的三个基色光的量规定为 R、G、B 三个基色单位，分别用[R]、[G]、[B]表示，当各以 1 个单位的上述红、绿、蓝三基色光混合时，恰能产生出等能白色（即 E 光源）。为了讨论方便，用 F 表示某一具有特定亮

度和色度的彩色光，于是 $F_{\text{白}}$ 的色匹配关系可用颜色方程式表示为：

$$F_{\text{白}}=1[\text{R}]+1[\text{G}]+1[\text{B}] \quad (2.4)$$

式中，等号的意义是“可由……混合配出”。

实验表明，匹配出等能白光三个基色单位[R]、[G]、[B]的光通量之比为 1:4.5907:0.0601。

这样，任意一种彩色光 F 可用下式表示：

$$F=\text{R}[\text{R}]+\text{G}[\text{G}]+\text{B}[\text{B}] \quad (2.5)$$

式中，R、G、B 称为红、绿、蓝三刺激值；R[R]、G[G]、B[B]称为 F 色光的三色分量。

1931 年 CIE 确定出混配等能单色辐射所需要的三刺激值，也就是说，为了混配出单位辐射功率、波长为 λ 的单色光所需要的三刺激值为光谱三刺激值，分别用 $\bar{r}(\lambda)$ 、 $\bar{g}(\lambda)$ 、 $\bar{b}(\lambda)$ 表示。于是单位辐射功率的三色光的颜色方程式为

$$F(\lambda)=\bar{r}(\lambda)[\text{R}]+\bar{g}(\lambda)[\text{G}]+\bar{b}(\lambda)[\text{B}] \quad (2.6)$$

我们对式 (2.3.2) 进行归一化处理，设 $m=\text{R}+\text{G}+\text{B}$ ，有

$$\begin{aligned} r &= \frac{\text{R}}{m} \\ g &= \frac{\text{G}}{m} \\ b &= \frac{\text{B}}{m} \\ r+g+b &= 1 \end{aligned} \quad (2.7)$$

式中， m 为色模，它代表某彩色光所含三基色的总量； r 、 g 、 b 称为色度坐标，它们分别表示当所用的三基色单位总量为 1 时，为配出某给定色度的彩色光所需要的 [R]、[G]、[B] 数值。将式 (2.7) 代入式 (2.5)，得到

$$F=m\{r[\text{R}]+g[\text{G}]+b[\text{B}]\} \quad (2.8)$$

对光谱三刺激值进行归一化处理，就可求出所有单色光的色度坐标：

$$\begin{aligned} r(\lambda) &= \bar{r}(\lambda)/[\bar{r}(\lambda)+\bar{g}(\lambda)+\bar{b}(\lambda)] \\ g(\lambda) &= \bar{g}(\lambda)/[\bar{r}(\lambda)+\bar{g}(\lambda)+\bar{b}(\lambda)] \\ b(\lambda) &= \bar{b}(\lambda)/[\bar{r}(\lambda)+\bar{g}(\lambda)+\bar{b}(\lambda)] \end{aligned} \quad (2.9)$$

不同波长单色光的光谱三刺激值相应特性和色度坐标如表 2.5 所示。

表 2.5 1931 CIE-RGB 色度系统的光谱三刺激值与色度坐标

光谱三刺激值相应特性			波长 / nm	色度坐标		
$\bar{r}(\lambda)$	$\bar{g}(\lambda)$	$\bar{b}(\lambda)$		$r(\lambda)$	$g(\lambda)$	$b(\lambda)$
0.000 03	-0.000 01	0.001 17	380	0.027 20	-0.011 50	0.984 30
0.000 10	-0.000 04	0.003 59	390	0.026 30	-0.011 40	0.985 10
0.000 30	-0.000 14	0.012 14	400	0.024 70	-0.011 20	0.986 50

续表

光谱三刺激值相应特性			波长 / nm	色度坐标		
$\bar{r}(\lambda)$	$\bar{g}(\lambda)$	$\bar{b}(\lambda)$		$r(\lambda)$	$g(\lambda)$	$b(\lambda)$
0.000 84	-0.000 41	0.037 07	410	0.022 50	-0.010 90	0.988 40
0.002 11	-0.001 10	0.115 41	420	0.018 10	-0.009 40	0.991 30
0.002 18	-0.001 19	0.247 69	430	0.008 80	-0.004 80	0.996 00
-0.002 61	0.001 49	0.312 28	440	-0.008 40	0.004 80	1.003 60
-0.012 13	0.006 78	0.316 70	450	-0.039 00	0.021 80	1.017 20
-0.026 08	0.014 85	0.298 21	460	-0.090 90	0.051 70	1.039 20
-0.039 33	0.025 38	0.229 91	470	-0.182 10	0.117 50	1.064 60
-0.049 39	0.039 14	0.144 94	480	-0.366 70	0.290 90	1.075 80
-0.058 14	0.056 98	0.082 57	490	-0.715 00	0.699 60	1.015 40
-0.071 73	0.085 36	0.047 76	500	-1.016 85	1.390 50	0.626 35
-0.089 01	0.128 60	0.026 98	510	-1.337 10	1.931 80	0.405 30
-0.092 64	0.174 68	0.012 21	520	-0.983 00	1.853 40	0.129 60
-0.071 01	0.203 17	0.005 49	530	-0.515 90	1.476 10	0.039 80
-0.031 52	0.214 66	0.001 46	540	-0.170 70	1.162 80	0.007 90
0.022 79	0.211 78	-0.000 58	550	0.097 40	0.905 10	-0.002 50
0.090 60	0.197 02	-0.001 30	560	0.316 40	0.688 10	-0.004 50
0.167 68	0.170 87	-0.001 35	570	0.497 30	0.506 70	-0.004 00
0.245 26	0.136 10	-0.001 08	580	0.644 90	0.357 90	-0.002 80
0.309 28	0.097 54	-0.000 79	590	0.761 70	0.240 20	-0.001 90
0.344 29	0.062 46	-0.000 49	600	0.847 50	0.153 70	-0.001 20
0.339 71	0.035 57	-0.000 30	610	0.905 90	0.094 90	-0.000 80
0.297 08	0.018 28	-0.000 15	620	0.942 50	0.058 00	-0.000 50
0.226 77	0.008 33	-0.000 08	630	0.964 90	0.035 40	-0.000 30
0.156 98	0.003 34	-0.000 03	640	0.979 70	0.020 50	-0.00020
0.101 67	0.001 16	-0.000 01	650	0.988 80	0.011 30	-0.000 10
0.059 32	0.000 37	0.000 00	660	0.994 00	0.006 10	-0.000 10
0.031 49	0.000 11	0.000 00	670	0.996 60	0.003 50	-0.000 10
0.016 87	0.000 03	0.000 00	680	0.998 40	0.000 16	0.001 44
0.008 19	0.000 00	0.000 00	690	0.999 60	0.000 40	0.000 00
0.004 10	0.000 00	0.000 00	700	1.000 00	0.000 00	0.000 00
0.002 10	0.000 00	0.000 00	710	1.000 00	0.000 00	0.000 00
0.001 05	0.000 00	0.000 00	720	1.000 00	0.000 00	0.000 00
0.000 52	0.000 00	0.000 00	730	1.000 00	0.000 00	0.000 00
0.000 25	0.000 00	0.000 00	740	1.000 00	0.000 00	0.000 00
0.000 12	0.000 00	0.000 00	750	1.000 00	0.000 00	0.000 00
0.000 06	0.000 00	0.000 00	760	1.000 00	0.000 00	0.000 00
0.000 03	0.000 00	0.000 00	770	1.000 00	0.000 00	0.000 00
0.000 00	0.000 00	0.000 00	780	1.000 00	0.000 00	0.000 00

由于三个色度坐标之和等于 1，因此只要用其中的两个坐标就可以明确地表示出色度，根据表 2.3.1 所列的各种谱色光的色度坐标值，可以绘出 380~780 nm 可见光谱在 r - g 直角坐标系中的位置，把它们逐点连线而形成一条舌形曲线，称为 1931 CIE- RGB 色度图，如图 2.9 所示。

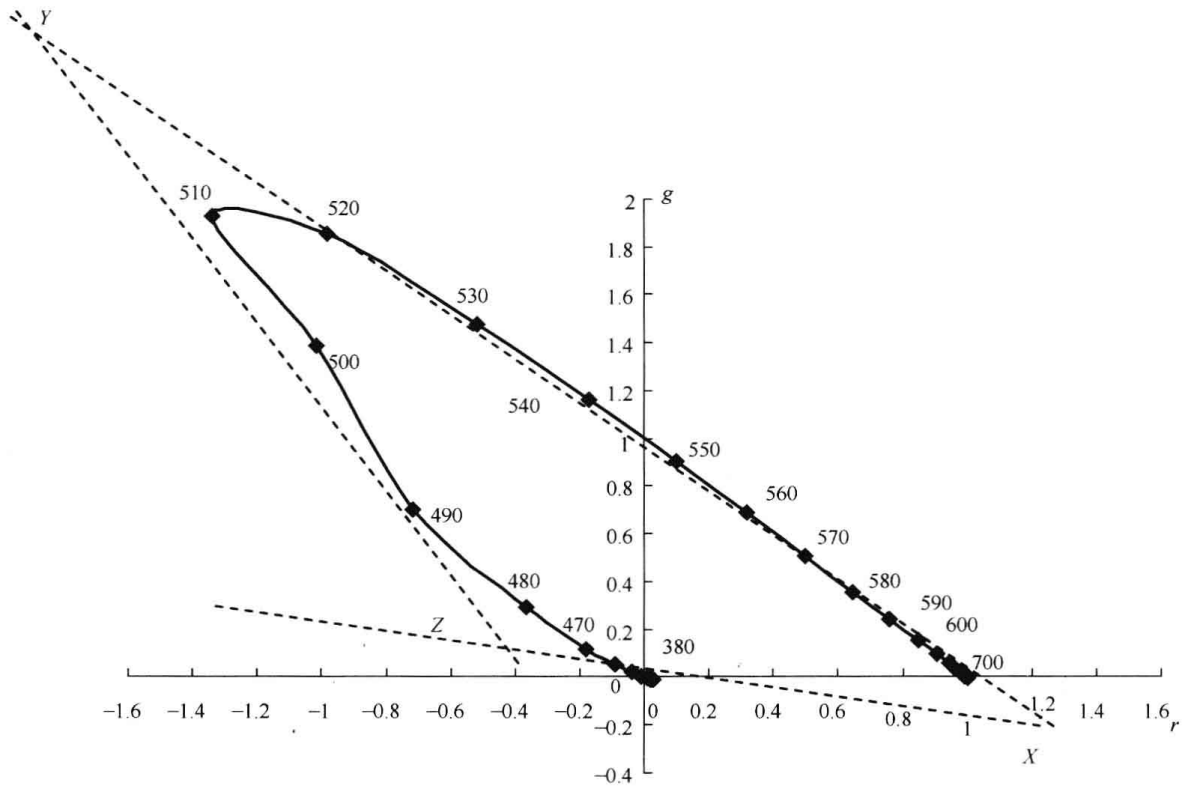


图 2.9 1931 CIE- RGB 色度图

在 1931 CIE- RGB 色度图中，单位红基色[R]的色度坐标为 $r=1$ 、 $g=b=0$ ；单位绿基色[G]的色度坐标为 $g=1$ 、 $r=b=0$ ；单位蓝基色[B]的色度坐标为 $r=g=0$ 、 $b=1$ ；当 $r=g=b=1/3$ 时，表示等能白光 $E_{白}$ 。由[R]、[G]、[B]三点连成的三角形称为 1931 CIE- RGB 彩色三角形，其质心 E 即为等能白光 $E_{白}$ 的位置。

自然界的彩色都能用舌形曲线及其内部的相应点坐标表示。坐标位置靠近谱色轨迹，彩色越纯，色饱和度越高；而越靠近 $E_{白}$ 点，彩色饱和度降低。

2.3.3 1931 CIE-XYZ 计色系统

由于 1931 CIE- RGB 色度图出现负坐标，使计算很不方便，CIE 在 1931 CIE- RGB 色度图的基础上，改用三个假想的三基色[X]、[Y]、[Z]建立了一个新的计色系统—1931 CIE-XYZ 计色系统。

在 1931 CIE-XYZ 计色系统中，假想的三基色[X]、[Y]、[Z]构成的三角形如

图 2.9 虚线三角形所示, 包含了 1931 CIE-RGB 色度图中的全部光谱轨迹。因此, 在这个新系统中, 光谱轨迹上以及光谱轨迹之内的颜色的色度坐标均为正值。

与 1931 CIE-RGB 计色系统相类似, 在 1931 CIE-XYZ 计色系统中, 用于匹配等能光谱刺激的 X、Y、Z 三基色数量称为光谱三刺激值, 分别用 $\bar{x}(\lambda)$ 、 $\bar{y}(\lambda)$ 、 $\bar{z}(\lambda)$ 表示。两种计色系统的三刺激值关系式为:

$$\begin{aligned}\bar{x}(\lambda) &= 2.7689 \bar{r}(\lambda) + 1.7518 \bar{g}(\lambda) + 1.1302 \bar{b}(\lambda) \\ \bar{y}(\lambda) &= 1.0000 \bar{r}(\lambda) + 4.5907 \bar{g}(\lambda) + 0.0601 \bar{b}(\lambda) \\ \bar{z}(\lambda) &= 0.0000 \bar{r}(\lambda) + 0.0565 \bar{g}(\lambda) + 5.5943 \bar{b}(\lambda)\end{aligned}\tag{2.10}$$

在 1931 CIE-XYZ 计色系统中, 规定 $\bar{y}(\lambda)$ 与明视觉光谱光效率函数一致, 即 $\bar{y}(\lambda) = V(\lambda)$ 。由光谱三刺激值 $\bar{x}(\lambda)$ 、 $\bar{y}(\lambda)$ 和 $\bar{z}(\lambda)$, 1931 CIE-XYZ 计色系统的色度坐标为 $x(\lambda)$ 、 $y(\lambda)$ 、 $z(\lambda)$:

$$\begin{aligned}x(\lambda) &= \bar{x}(\lambda) / [\bar{x}(\lambda) + \bar{y}(\lambda) + \bar{z}(\lambda)] \\ y(\lambda) &= \bar{y}(\lambda) / [\bar{x}(\lambda) + \bar{y}(\lambda) + \bar{z}(\lambda)] \\ z(\lambda) &= \bar{z}(\lambda) / [\bar{x}(\lambda) + \bar{y}(\lambda) + \bar{z}(\lambda)]\end{aligned}\tag{2.11}$$

根据式 (2.10) 和式 (2.11), 可以得出 CIE 1931-XYZ 光谱色度系统的光谱三刺激值与色度坐标, 如表 2.6 所示。

表 2.6 1931 CIE-XYZ 色度系统的光谱三刺激值与色度坐标

光谱三刺激值相应特性			波长 / nm	色度坐标		
$\bar{x}(\lambda)$	$\bar{y}(\lambda)$	$\bar{z}(\lambda)$		$x(\lambda)$	$y(\lambda)$	$z(\lambda)$
0.001 4	0.000 1	0.006 5	380	0.173 8	0.006 8	0.819 4
0.004 3	0.000 1	0.020 1	390	0.174 2	0.005 4	0.820 4
0.014 3	0.000 4	0.067 9	400	0.173 2	0.004 7	0.822 1
0.043 5	0.001 2	0.207 4	410	0.172 6	0.004 7	0.822 7
0.134 4	0.004 0	0.645 6	420	0.171 4	0.005 1	0.823 5
0.283 9	0.011 6	1.385 6	430	0.168 9	0.006 9	0.824 2
0.348 3	0.023 0	1.747 1	440	0.164 4	0.010 9	0.824 7
0.336 2	0.038 0	1.772 1	450	0.156 6	0.017 7	0.825 6
0.290 8	0.060 0	1.669 1	460	0.144 0	0.029 7	0.826 3
0.195 4	0.091 0	1.287 6	470	0.124 1	0.057 8	0.818 0
0.095 6	0.139 0	0.813 0	480	0.091 3	0.132 7	0.776 1
0.032 2	0.208 4	0.465 1	490	0.045 6	0.295 3	0.659 1
0.004 9	0.323 0	0.272 0	500	0.008 2	0.538 4	0.453 4
0.009 3	0.503 0	0.158 2	510	0.013 9	0.750 2	0.235 9
0.063 3	0.710 0	0.078 2	520	0.074 3	0.833 9	0.091 8
0.165 5	0.862 0	0.042 2	530	0.154 7	0.805 8	0.039 4
0.290 4	0.954 0	0.020 3	540	0.229 6	0.754 3	0.016 0

续表

光谱三刺激值相应特性			波长 / nm	色度坐标		
$\bar{x}(\lambda)$	$\bar{y}(\lambda)$	$\bar{z}(\lambda)$		$x(\lambda)$	$y(\lambda)$	$z(\lambda)$
0.433 4	0.995 0	0.008 7	550	0.301 6	0.692 3	0.006 1
0.594 5	0.995 0	0.003 9	560	0.373 1	0.624 4	0.002 4
0.762 1	0.952 0	0.002 1	570	0.444 1	0.554 7	0.001 2
0.916 3	0.870 0	0.001 6	580	0.512 5	0.486 6	0.000 9
1.026 3	0.757 0	0.001 1	590	0.575 2	0.424 2	0.000 6
1.062 2	0.631 0	0.000 8	600	0.627 0	0.372 5	0.000 5
1.002 6	0.503 0	0.000 3	610	0.665 8	0.334 0	0.000 2
0.854 4	0.381 0	0.000 2	620	0.691 5	0.308 3	0.000 2
0.642 4	0.265 0	0.000 0	630	0.707 9	0.292 0	0.000 0
0.440 5	0.172 3	0.000 0	640	0.718 8	0.281 2	0.000 0
0.283 5	0.107 0	0.000 0	650	0.726 0	0.274 0	0.000 0
0.164 9	0.061 0	0.000 0	660	0.729 8	0.270 1	0.000 1
0.087 4	0.032 0	0.000 0	670	0.732 0	0.268 0	0.000 1
0.046 8	0.017 0	0.000 0	680	0.733 3	0.266 7	0.000 0
0.022 7	0.008 2	0.000 0	690	0.734 7	0.265 3	0.000 0
0.011 4	0.004 1	0.000 0	700	0.734 7	0.265 3	0.000 0
0.005 8	0.002 1	0.000 0	710	0.734 7	0.265 3	0.000 0
0.002 9	0.001 1	0.000 0	720	0.734 7	0.265 3	0.000 0
0.001 4	0.000 5	0.000 0	730	0.734 7	0.265 3	0.000 0
0.000 7	0.000 3	0.000 0	740	0.734 7	0.265 3	0.000 0
0.000 3	0.000 1	0.000 0	750	0.734 7	0.265 3	0.000 0
0.000 2	0.000 1	0.000 0	760	0.734 7	0.265 3	0.000 0
0.000 1	0.000 0	0.000 0	770	0.734 7	0.265 3	0.000 0
0.000 0	0.000 0	0.000 0	780	0.734 7	0.265 3	0.000 0

$x(\lambda)$ 、 $y(\lambda)$ 、 $z(\lambda)$ 是归一化值， $x(\lambda)+y(\lambda)+z(\lambda)=1$ 。在三个色度坐标中，有一个是独立的，相应地使用 x - y 坐标来表示各种色度，这就是人们常说的舌形色度图，如图 2.10 所示。

在 1931 CIE-XYZ 色度图中，舌形曲线全部位于第一象限，所有可见光的光谱色都在封闭曲线内，所有的单色光都位于舌形曲线上。

舌形图可以表示任意颜色的三要素： y 轴的大小代表亮度；从等能量 E 白光的坐标 W_E 到谱色曲线的任一点连线称为等色调线。该线上所有的点均对应同一色调的彩色；线上的点离 W_E 越近，该点对应的色饱和度就越小。

在舌形曲线内任取三点对应的彩色作基色，则由此三基色混合成的所有彩色都包含在以这三点为顶点的三角形内。三角形外的彩色不能由此三基色混得到。

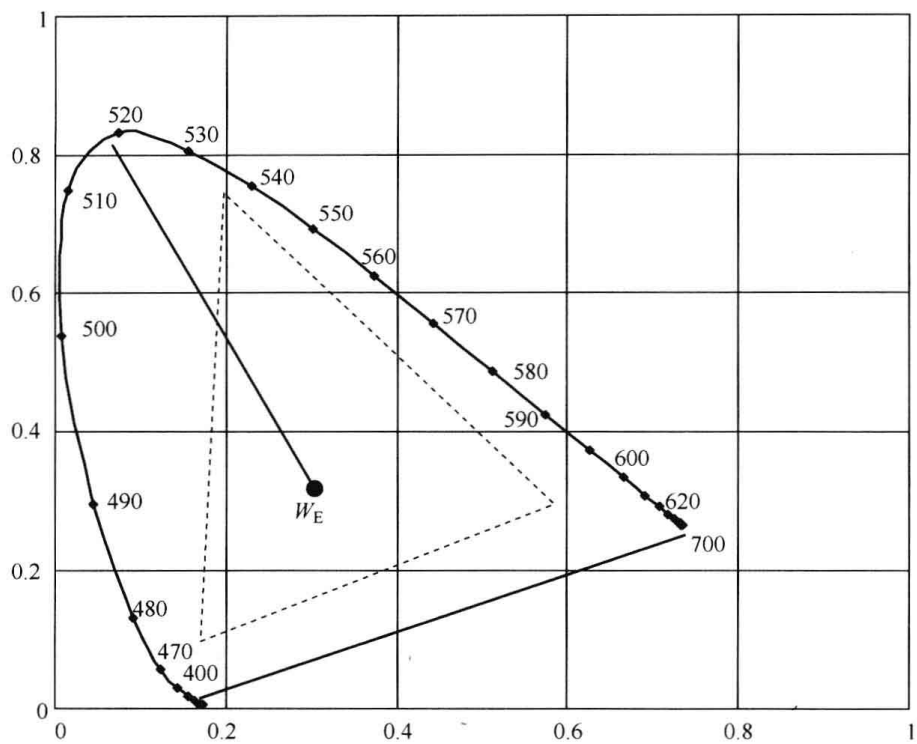


图 2.10 1931 CIE-XYZ 色度图

2.3.4 1960 CIE-UCS 计色系统

根据 2.2.2 节所述，人眼对不同波长的光有不同的色调感觉，例如，人眼对 480~640 nm 区间的光的色调分辨力较高，其中对 500 nm（青绿）和 600 nm（橙黄）两个波长来说，只要波长变化 1 nm，便可以分辨出颜色色调的变化；而有的波长区间，例如，655 nm 的红色到可见光谱长波的长端，以及 430 nm 的紫色到可见光谱的短端，人眼几乎看不到色调的差别。

因此，在 1931 CIE-XYZ 系统的色度图上，每一种颜色虽然对应一个点，但对人眼来说却是一个范围，也就是说色度图上的位置发生了变化，人眼不一定能发现其中的变化。人眼感觉不出颜色的变化的范围叫做颜色的宽容度。正常视力的观察者能看出颜色变化的最小距离称为颜色的刚辨差（Just Noticeable Difference, JND）。

由于 1931 CIE-XYZ 系统的色度的色度空间在视觉上时不均匀的，1960 年 CIE 制定了 1960 CIE 均匀色度标尺图（1960 CIE Uniform Chromaticity-Scale Diagram），简称 1960 CIE-UCS 色度图。在 1960 CIE-UCS 色度图上，颜色的宽容度轨迹近似于圆形。

1960 CIE-UCS 色度系统的三刺激值为 $\bar{u}(\lambda)$ 、 $\bar{v}(\lambda)$ 、 $\bar{w}(\lambda)$ ，它与 1931 CIE-XYZ 计色系统的三刺激值 $\bar{x}(\lambda)$ 、 $\bar{y}(\lambda)$ 、 $\bar{z}(\lambda)$ 关系为

$$\begin{aligned} \bar{u}(\lambda) &= \frac{2}{3} \bar{x}(\lambda) \\ \bar{v}(\lambda) &= \bar{y}(\lambda) \\ \bar{w}(\lambda) &= \frac{1}{2} [-\bar{x}(\lambda) + 3\bar{y}(\lambda) + \bar{z}(\lambda)] \end{aligned} \tag{2.12}$$

其归一化值为 1960 CIE-UCS 色度系统的色度坐标值:

$$\begin{aligned} u(\lambda) &= \bar{u}(\lambda) / \{ \bar{u}(\lambda) + \bar{v}(\lambda) + \bar{w}(\lambda) \} \\ v(\lambda) &= \bar{v}(\lambda) / \{ \bar{u}(\lambda) + \bar{v}(\lambda) + \bar{w}(\lambda) \} \\ w(\lambda) &= \bar{w}(\lambda) / \{ \bar{u}(\lambda) + \bar{v}(\lambda) + \bar{w}(\lambda) \} \end{aligned} \tag{2.13}$$

利用表 2.6 及式 (2.12)、式 (2.13)，可以得到 160 CIE-UCS 色度系统的光谱三刺激值与色度坐标，如表 2.7 所示。图 2.11 表示出经过以上变换后得到的 1960 CIE-UCS 色度图。

表 2.7 1960 CIE-UCS 色度系统的光谱三刺激值与色度坐标

光谱三刺激值相应特性			波长 / nm	色度坐标		
$\bar{u}(\lambda)$	$\bar{v}(\lambda)$	$\bar{w}(\lambda)$		$u(\lambda)$	$v(\lambda)$	$w(\lambda)$
0.000 9	0.000 1	0.002 7	380	0.254 2	0.014 9	0.730 8
0.002 8	0.000 1	0.008 1	390	0.256 5	0.011 9	0.731 5
0.009 5	0.000 4	0.027 4	400	0.255 7	0.010 4	0.734 0
0.029 0	0.001 2	0.083 7	410	0.254 6	0.010 4	0.734 9
0.089 6	0.004 0	0.261 6	420	0.252 2	0.011 3	0.736 6
0.189 3	0.011 6	0.568 3	430	0.246 1	0.015 1	0.738 8
0.232 2	0.023 0	0.733 9	440	0.234 8	0.023 3	0.742 0
0.224 1	0.038 0	0.775 0	450	0.216 1	0.036 7	0.747 2
0.193 9	0.060 0	0.779 2	460	0.187 7	0.058 1	0.754 2
0.130 3	0.091 0	0.682 6	470	0.144 1	0.100 7	0.755 2
0.063 7	0.139 0	0.567 2	480	0.082 8	0.180 5	0.736 7
0.021 4	0.208 4	0.529 1	490	0.028 2	0.274 6	0.697 2
0.003 3	0.323 0	0.618 1	500	0.003 5	0.342 0	0.654 5
0.006 2	0.503 0	0.828 9	510	0.004 6	0.375 9	0.619 5
0.042 2	0.710 0	1.072 4	520	0.023 1	0.389 1	0.587 8
0.110 3	0.862 0	1.231 4	530	0.050 1	0.391 2	0.558 8
0.193 6	0.954 0	1.296 0	540	0.079 2	0.390 4	0.530 4
0.289 0	0.995 0	1.280 1	550	0.112 7	0.388 0	0.499 3
0.396 4	0.995 0	1.197 1	560	0.153 1	0.384 4	0.462 5
0.508 1	0.952 0	1.048 0	570	0.202 6	0.379 6	0.417 9
0.610 9	0.870 0	0.847 7	580	0.262 3	0.373 6	0.364 0
0.684 2	0.757 0	0.622 9	590	0.331 5	0.366 7	0.301 8
0.708 1	0.631 0	0.415 8	600	0.403 5	0.359 6	0.236 9

续表

光谱三刺激值相应特性			波长 / nm	色度坐标		
$\bar{u}(\lambda)$	$\bar{v}(\lambda)$	$\bar{w}(\lambda)$		$u(\lambda)$	$v(\lambda)$	$w(\lambda)$
0.668 4	0.503 0	0.253 3	610	0.469 1	0.353 0	0.177 8
0.569 6	0.381 0	0.144 4	620	0.520 2	0.347 9	0.131 8
0.428 3	0.265 0	0.076 3	630	0.556 5	0.344 3	0.099 2
0.293 7	0.172 3	0.038 2	640	0.582 4	0.341 8	0.075 8
0.189 0	0.107 0	0.018 7	650	0.600 6	0.339 9	0.059 5
0.109 9	0.061 0	0.009 1	660	0.610 6	0.338 9	0.050 5
0.058 3	0.032 0	0.004 3	670	0.616 1	0.338 4	0.045 5
0.031 2	0.017 0	0.002 1	680	0.619 6	0.338 0	0.042 3
0.015 1	0.008 2	0.000 9	690	0.623 3	0.337 7	0.039 0
0.007 6	0.004 1	0.000 5	700	0.623 3	0.337 7	0.039 0
0.003 9	0.002 1	0.000 2	710	0.623 3	0.337 7	0.039 0
0.001 9	0.001 1	0.000 1	720	0.623 3	0.337 7	0.039 0
0.001 0	0.000 5	0.000 1	730	0.623 3	0.337 7	0.039 0
0.000 5	0.000 3	0.000 0	740	0.623 3	0.337 7	0.039 0
0.000 2	0.000 1	0.000 0	750	0.623 3	0.337 7	0.039 0
0.000 1	0.000 1	0.000 0	760	0.623 3	0.337 7	0.039 0
0.000 1	0.000 0	0.000 0	770	0.623 3	0.337 7	0.039 0
0.000 0	0.000 0	0.000 0	780	0.623 3	0.337 7	0.039 0

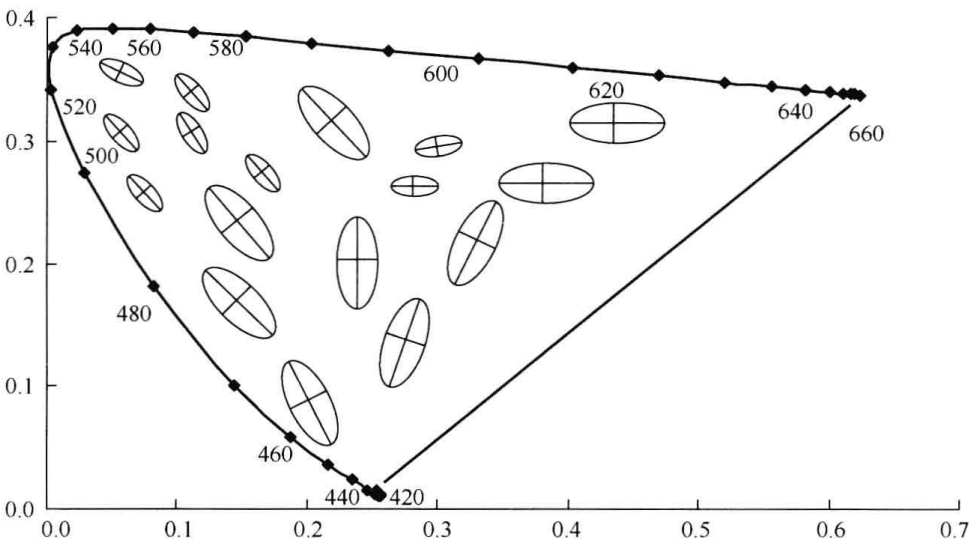


图 2.11 1960 CIE-UCS 色度图

在 1960 CIE-UCS 色度图上用下列公式计算刚辨差：

$$\Delta C_{JND} = \frac{\sqrt{(u_0 - u)^2 + (v_0 - v)^2}}{0.00384} \text{ (JND)} \tag{2.14}$$

式中， u_0 和 v_0 是特定点的色度坐标； u 和 v 是偏离特定点的色度坐标。

2.3.5 1964 CIE-UCS 均匀颜色空间 (W^* 、 U^* 、 V^*)

1960 CIE-UCS 色度图没有给出明度坐标,所以在给出 u 、 v 坐标时还必须单独给出 Y 值。由于实际应用的许多颜色问题均与物体的亮度有关,因此 1964 年 CIE 规定了“均匀颜色空间”,将 1960 CIE-UCS 色度图的二维空间扩充到包括亮度因数在内的三维均匀空间。在均匀颜色空间中,色差计算可不限于具有相等的亮度因数的颜色。1964 CIE 均匀颜色空间用明度指数 W^* 和色度指数 U^* 、 V^* 坐标系统来表示。 W^* 、 U^* 、 V^* 坐标是根据三刺激值规定的:

$$\begin{aligned} W^* &= 25 Y^{\frac{1}{3}} - 17 \quad 1 \leq Y \leq 100 \\ U^* &= 13 W^* (u - u_0) \\ V^* &= 13 W^* (v - v_0) \end{aligned} \quad (2.15)$$

式中, Y 为亮度值, u 、 v 为颜色的色度坐标; u_0 、 v_0 为参考标准光源的色度坐标。

u 、 v 与 1960 CIE-UCS 色度图坐标 x 、 y 的关系为

$$\begin{aligned} u &= \frac{4x}{-2x + 12y + 3} \\ v &= \frac{6y}{-2x + 12y + 3} \end{aligned} \quad (2.16)$$

用 1964 均匀颜色空间的三维空间概念,可以通过公式计算两个颜色 W_1^* 、 U_1^* 、 V_1^* 和 W_2^* 、 U_2^* 、 V_2^* 之间的差异 ΔE , 计算公式如下:

$$\Delta E = \sqrt{\Delta U^{*2} + \Delta V^{*2} + \Delta W^{*2}} \quad (2.17)$$

当 $\Delta E = 1$ 时,称为 1 个 NBS (National Bureau of Standards, 美国国家标准局) 色差单位,1 个 NBS 色差单位大约相当于在最优实验条件下人眼所能察觉的刚辨差 (JND) 的 5 倍。

2.3.6 1976 CIE- L^* u^* v^* 均匀颜色空间

为了进一步统一评价颜色差别的方法, CIE 于 1976 年又提出了 1976 CIE- L^* u^* v^* 颜色空间,其表达式为

$$\begin{aligned} L^* &= 116 \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - 16 \quad \frac{Y}{Y_0} > 0.01 \\ u^* &= 13 L^* (u' - u'_0) \\ v^* &= 13 L^* (v' - v'_0) \end{aligned} \quad (2.18)$$

式中, u' 和 v' 为颜色的色度坐标, u'_0 和 v'_0 为参考标准光源的色度坐标, $Y_0 = 100$ 。

1976 CIE- L^* u^* v^* 与 1931 CIE-XYZ 色度空间的坐标关系与式 (2.16) 相同。

按 1976 CIE- $L^*u^*v^*$ 颜色空间标定的两个颜色的总色差 $\Delta E(L^*u^*v^*)$ 为:

$$\Delta E(L^*u^*v^*) = \sqrt{\Delta L^{*2} + \Delta u^{*2} + \Delta v^{*2}} \quad (2.19)$$

1976 CIE- $L^*u^*v^*$ 与 1964 CIE- $L^*U^*V^*$ 的主要区别在于改变了 $u-v$ 色度图的 v 坐标, 即 $v'=1.5v$, 而这两种颜色空间的 u 坐标保持不变, 即 $u'=u$ 。

两个国际标准化组织 IEC 和 ITU 在规范视频设备 (包括数字电视显示器、电视接收机、各种投影电视机等) 的色度误差时, 推荐使用 1976 CIE $u'-v'$ 色度系统, 在我国数字电视显示器 (LCD、PDP) 行业标准中, 用它规范白场相关色温、色域覆盖率、白色色度不均匀性、白平衡误差、色度与视觉的关系等参数, 后续章节将会涉及相关内容。

第3章 现代电视技术基础

现代视频会议的多媒体技术基础无疑与广播电视相关领域的技术和设备密切相关，本章将重点介绍下列现代电视的基础知识，作为视频会议系统相关设备和技术论述的一个前提铺垫。

本章知识要点如下：

- 模拟电视技术基础；
- 数字电视技术基础；
- 主要视频接口特性综述。

3.1 模拟电视技术基础

3.1.1 逐行扫描与隔行扫描

在早期的模拟电视中，电视图像的显示是通过电子束轰击涂在玻璃屏上的荧光粉来实现的。

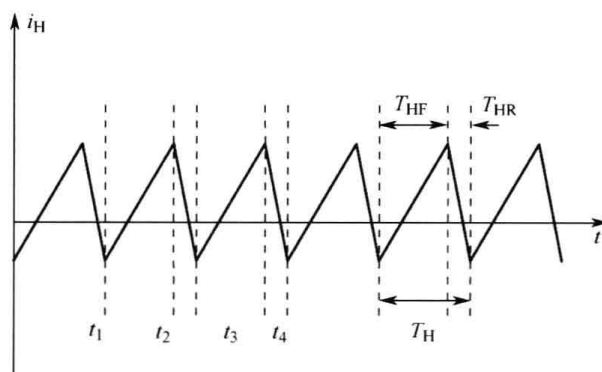
因此，只要用代表图像模拟信号区控制电子束按照规定的次序扫描荧光屏，便可以重显电视图像。显像管中的电子束是通过水平和垂直偏转线圈来实现的。在水平偏转线圈所产生的垂直磁场的作用下，电子束沿水平方向扫描，称之为行扫描。在垂直偏转线圈所产生的水平磁场作用下，电子束沿垂直方向扫描，称之为帧扫描。

假定在水平偏转线圈通过如图 3.1.1 (a) 所示的锯齿形电流，当 $t_1 \sim t_2$ 期间的电流线性增长时，电子束在垂直偏转磁场的作用下从左向右作匀速扫描叫作行扫描正程，当 t_2 时刻正程结束时，电子束扫到屏幕的最右边。当 $t_2 \sim t_3$ 期间的电流快速线性减小时，电子束从右向左作迅速扫描，叫作行扫描逆程，当 t_3 时刻逆程结束时，电子束又回到屏幕的最左边。电子束在水平方向往返一次所需要的时间称为行扫描周期 T_H 。 T_H 等于行正程时间 T_{HF} 与行逆程时间 T_{HR} 之和，行扫描周期的倒数就是行扫描频率 f_H 。

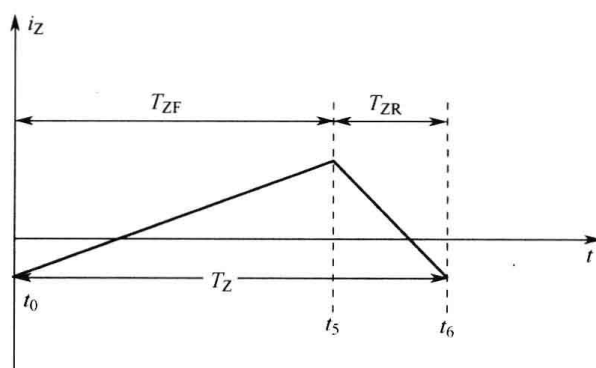
同理，在垂直偏转线圈加上如图 3.1.1 (b) 所示的锯齿形电流，将产生帧扫描过程。电子束在垂直方向往返一次所需要的时间称为帧扫描周期 T_Z 。 T_H 等于帧正

程时间 T_{ZF} 与帧逆程时间 T_{ZR} 之和，帧扫描周期的倒数就是帧扫描频率 f_Z 。帧扫描频率 f_Z 远小于行扫描频率 f_H 。

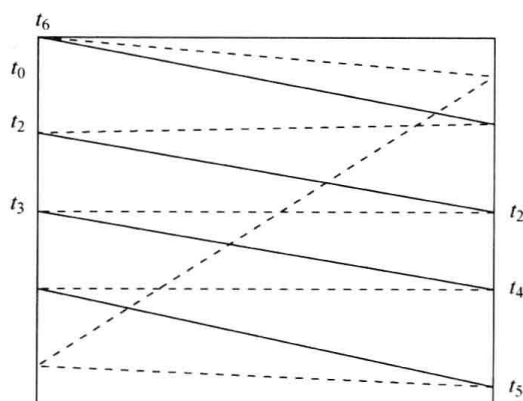
如果把上述两种电流同时加到水平和垂直偏转线圈上，电子束将同时沿水平和垂直方向扫描。在屏幕上就显示出如图 3.1 (c) 所示的光栅。由于帧扫描频率 f_Z 远小于行扫描频率 f_H ，因此电视图像的扫描线看起来是水平直线。这种从图像的上端开始、从左到右，从上到下以均匀速度按次序一行接一行地扫完完全的帧画面，称为逐行扫描。



(a) 行扫描电流波形



(b) 帧扫描电流波形



(c) 扫描光栅

图 3.1 逐行扫描

在逐行扫描中,所有帧的光栅都应相互重合,这就要求帧扫描周期 T_Z 是行扫描周期 T_H 的整数倍,也就是说每帧的扫描行数为整数,即 $T_Z = ZT_H$ 。

为了节省电视信号带宽,同时又保证图像的清晰度。把一帧图像分成两场来扫描,第一场扫描 1、3、5、…奇数行,称为奇数场。第二场扫描 2、4、6、…偶数行,称为偶数场。这样,每帧图像经过两场扫描,所有像素全部扫完。

按照隔行扫描的需要,每帧图像的最大标称扫描行数 Z 应为奇数。在隔行扫描电视系统中,每场图像的扫描行数应为 $(n + \frac{1}{2})$ 行,奇数场从第一行开始,结束于半行,偶数场从半行开始结束于整数行,奇数场与偶数场相嵌,构成一个完整帧,如图 3.2 所示。

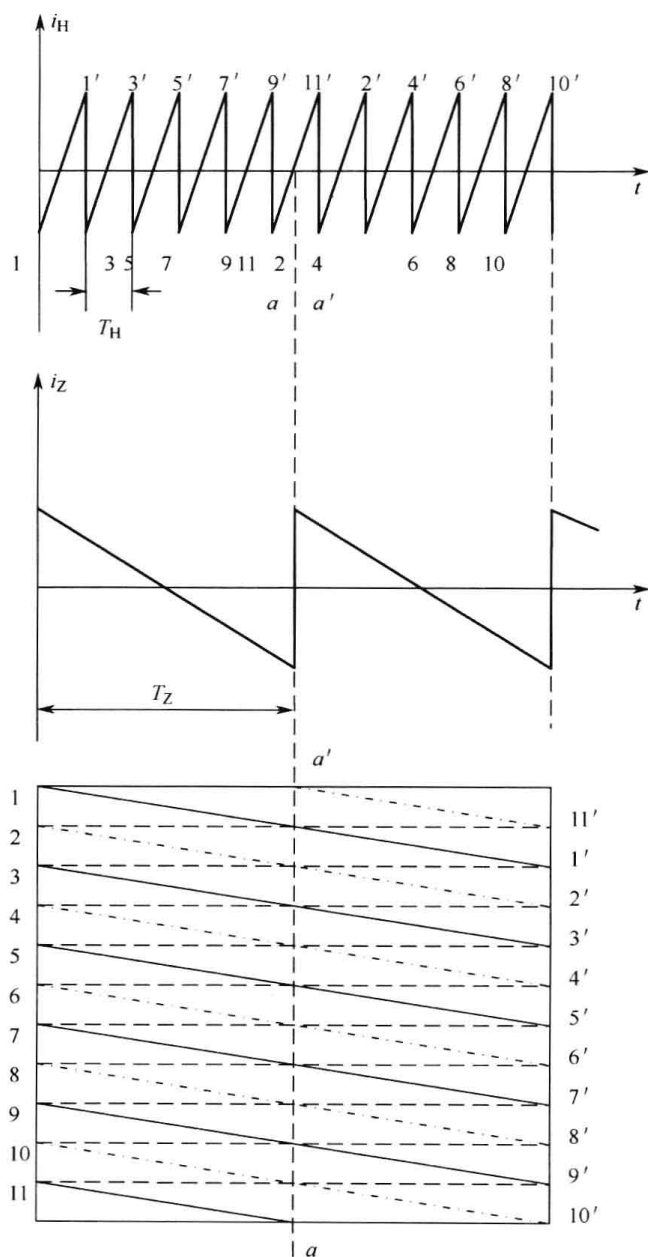


图 3.2 隔行扫描

我国模拟电视和标清电视标准规定：每帧图像的扫描行数 Z 为 625 行，奇数场和偶数场个占 312.5 行。由于存在 50 个消隐行，被显示的水平扫描行为 575 行。场频 $f_v=50$ Hz。由于是隔行扫描，所以帧频 $f_z=25$ Hz，则一帧的扫描时间 $T_z=40$ ms。当扫描行数 $Z=625$ 后，行扫描时间 $T_H=T_z/Z=40\text{ ms}/625=64\text{ us}$ ，行频 $f_H=f_z \times Z=25\text{ Hz} \times 625=15\,625\text{ Hz}$ 。

NTSC 制标准规定：每帧图像的扫描行数 Z 为 525 行，奇数场和偶数场个占 262.5 行。由于存在 42 个消隐行，被显示的水平扫描行为 483 行。场频 $f_v=59.94$ Hz。由于是隔行扫描，所以帧频 $f_z=29.97$ Hz，行频 $f_H=f_z \times Z=29.94\text{ Hz} \times 525=15\,734\text{ Hz}$ 。

高清电视标准规定：每帧图像的扫描行数 Z 为 1 125 行，奇数场和偶数场个占 562.5 行。被显示的水平扫描行为 1 080 行，场频为 50 Hz，行频为 28 125 Hz。

3.1.2 亮度方程式

目前，常见的 NTSC 制和 PAL 制彩色电视系统其选用的标准白光光源是不一样的，NTSC 制选用 C 光源，PAL 制选用 D₆₅ 光源，如表 2.2 所示。

NTSC 制规定的显像三基色 Re、Ge、Be 坐标转换为 1931 CIE-XYZ 坐标，其关系为

$$\begin{aligned} X &= 0.607 R_e + 0.174 G_e + 0.200 B_e \\ Y &= 0.299 R_e + 0.587 G_e + 0.114 B_e \\ Z &= 0.000 R_e + 0.066 G_e + 1.116 B_e \end{aligned} \quad (3.1)$$

式中， Y 代表彩色的亮度，则由显像三基色配出的任意彩色光的亮度为

$$Y = 0.299R + 0.587G + 0.114B \quad (3.2)$$

式 (3.2) 是彩色电视系统广泛使用的亮度方程式，它是电视中心设备、测试仪器、彩色电视机等进行编解码的主要计算公式。由于 NTSC 制式彩色电视的理论基础，发展历史最久，尽管 PAL 制选用的标准光源于 NTSC 制不同，相应的亮度方程式系数也不同，但也同样采用了 NTSC 制的亮度方程式。虽然理论上存在误差，但性能上仍满足视觉对亮度的要求。

根据亮度方程式 $Y=0.299R+0.587G+0.114B$ ，我们得到三个色差信号： $R-Y$ ， $B-Y$ ， $G-Y$ 。

$$\begin{aligned} R-Y &= 0.701R - 0.587G - 0.114B \\ B-Y &= -0.299R - 0.587G + 0.866B \\ G-Y &= -0.299R + 0.413G - 0.114B \end{aligned} \quad (3.3)$$

三个色差信号中只有两个是独立的，第三个可以由另外两个得到，因此只要选择两个色差信号就可以代表色度信号。由于 $G-Y$ 的信号信噪比较其他两者小，

且由其产生其他色差信号导致增加系统的复杂性, 于是所有的电视制式均选作 $R-Y$ 和 $B-Y$ 作为色差信号。

电视系统传送色差信号而不是色度信号的优势在于:

① 可以减少色度信号对亮度信号的干扰, 当传送黑白图像时, $R=G=B$, 两个色差信号 $R-Y$ 和 $B-Y$ 均为零, 不会对亮度信号产生干扰。

② 能够实现亮度恒定原理, 即重现图像的亮度只由传送亮度信息的亮度信号决定。

③ 可以节省色度信号的发射功率。在彩色图像中, 大部分像素接近于白色或灰色, 它们的色差信号为零, 小部分彩色像素才有色差信号, 因此发射色差信号比发射 R 、 G 、 B 信号需要的功率小。

3.1.3 NTSC 制电视信号

根据图像彩色三要素原则, 模拟电视系统发射一个亮度信号和两个色差信号供彩色电视重显彩色电视图像, 亮度信号单独供黑白电视重显黑白电视图像。

早期的电视传输系统的传输带宽为 $0\sim 6\text{ MHz}$, 首先在 $0\sim 6\text{ MHz}$ 中选择一个频率作为彩色副载波, 用两个色差信号对彩色副载波进行调制, 调制后的信号称为色度信号。色度信号、亮度信号、同步信号叠加后为彩色全电视信号。全电视信号再去调制图像载波, 经射频放大后发射出去。

设色差信号之一 $u_{R-Y}=(R-Y)\cos\Omega t$ 对副载波 $u_{SC}=U_{SC}\cos\omega_{sc}t$ 进行调幅, 则调幅后的信号数学表达式为

$$u_{AM}=U_{SC}\left\{\cos\omega_{sc}t+\frac{m}{2}\cos(\omega_{sc}-\Omega)t+\frac{m}{2}\cos(\omega_{sc}+\Omega)t\right\} \quad (3.4)$$

式中, $m=(R-Y)/U_{SC}$ 。

式 (3.4) 表明了调幅波包含了三个频率: 副载波频率 ω_{sc} 和两个边频 $\omega_{sc}\pm\Omega$ 。由于副载波 ω_{sc} 上不带任何信息, 所以把副载波抑制掉可以节省发射功率。副载波抑制掉称为平衡调幅波, 平衡调幅波的数学表达式为

$$\begin{aligned} u_{BM} &= U_{SC}\left\{\frac{m}{2}\cos(\omega_{sc}-\Omega)t+\frac{m}{2}\cos(\omega_{sc}+\Omega)t\right\}=m U_{SC}\cos\Omega t\times\cos\omega_{sc}t \\ &= (R-Y)\cos\Omega t\times\cos\omega_{sc}t \end{aligned} \quad (3.5)$$

式 (3.5) 表明, 用一个乘法器将色差信号与副载波信号相乘可以得到平衡调幅波。

为了在同一频带内传送两个色差信号 $R-Y$ 和 $B-Y$, 要将两个色差信号进行正交平衡调幅, 用两个色差信号 $R-Y$ 和 $B-Y$ 分别对频率相同、相位差 90° 的两个副载波 $\cos\omega_{sc}t$ 和 $\sin\omega_{sc}t$ 进行平衡调幅, 然后通过线性相加器形成色度信号, 如图 3.3 所示。

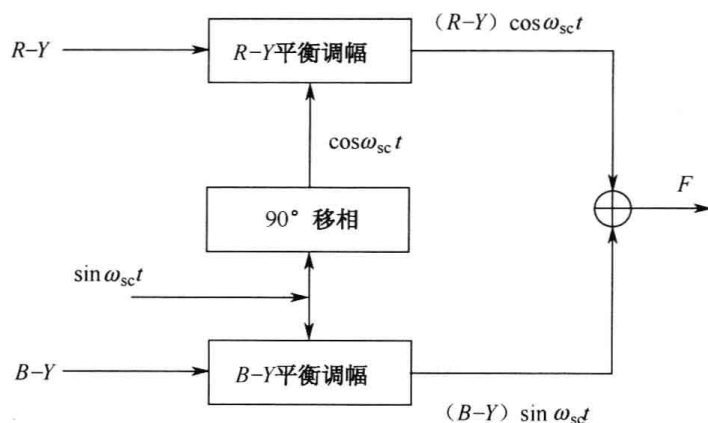


图 3.3 色度信号正交平衡调幅

设色度信号为 F ，根据图 3.3 的表达，色度信号 F 的数学表达式为

$$\begin{aligned}
 F &= (B-Y)\sin\omega_{sc}t + (R-Y)\cos\omega_{sc}t \\
 &= \sqrt{(B-Y)^2 + (R-Y)^2} \sin(\omega_{sc}t + \theta) \\
 &= F_m \sin(\omega_{sc}t + \theta)
 \end{aligned} \tag{3.6}$$

式中， $F_m = \sqrt{(B-Y)^2 + (R-Y)^2}$ ， $\theta = \arctan \frac{R-Y}{B-Y}$ 。

因此，色度信号的振幅和相角中包含了彩色图像的全部色度信息，振幅 F_m 取决于色差信号的幅值，决定了所传彩色的饱和度；而相角 θ 取决于色差信号的相对比值，决定了彩色的色调。也就是说，色度信号是一个既调幅又调相的波形，其幅值传达了图像的色饱和度，其相位传送了图像的色调。

根据式 (3.6) 画出了色度信号的矢量图，图 3.4 中对角线的长度代表色度信号的幅值，而 θ 为相角。

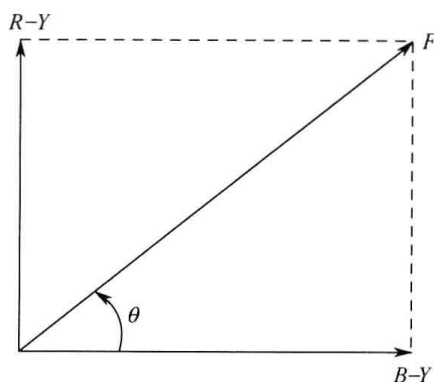


图 3.4 色度信号的矢量图

在正式调制全电视信号之前，需要对色度信号进行一定程度的压缩。压缩后的色差信号 $R-Y$ 用 V 表示，压缩后的色差信号 $B-Y$ 用 U 表示：

$$\begin{aligned}
 V &= 0.877(R-Y) \\
 U &= 0.493(B-Y)
 \end{aligned} \tag{3.7}$$

用 U 、 V 信号区调制副载波，这样彩色电视色度信号表达式为

$$\begin{aligned} F &= F_U + F_V = U \sin \omega_{sc} t + V \cos \omega_{sc} t \\ &= \sqrt{U^2 + V^2} \sin(\omega_{sc} t + \theta) \\ &= F_m \sin(\omega_{sc} t + \theta) \end{aligned} \quad (3.8)$$

式中， $\theta = \arctan \frac{V}{U}$ 。

NTSC 制色度信号采用正交平衡调幅调制方式，又称正交平衡调幅制。NTSC 制色度信号的组成方式最简单，解码电路也最简单。

电视系统会有各种各样的相位失真，其中最主要的传输系统非线性引起的微分相位失真。NTSC 制对相位失真比较敏感，容易产生色调畸变。一般相位失真超过 $\pm 5^\circ$ ，人眼就会觉察出色调的失真。

为了克服 NTSC 制的相位敏感性，又提出了交平衡调幅制，即 PAL 制（Phase Alternation Line）。

3.1.4 PAL 制电视信号

所谓逐行倒相，是将色度信号的一个分量，即 F_V 进行逐行倒相。PAL 制色度信号的数学表达式为

$$\begin{aligned} F &= F_U \pm F_V = U \sin \omega_{sc} t \pm V \cos \omega_{sc} t \\ &= \sqrt{U^2 + V^2} \sin(\omega_{sc} t \pm \theta) \\ &= F_m \sin(\omega_{sc} t \pm \theta) \end{aligned} \quad (3.9)$$

式中， $\theta = \arctan \frac{V}{U}$ 。

式 (3.9) 中的 $\pm$ 号表示：第 n 行（因为这一行与 NTSC 制一样，又称 NTSC 行）取正号，通常用矢量 F_n 表示；第 $n+1$ 行（又称 PAL 行）取负号，通常用矢量 F_{n+1} 表示。因为 F_n 和 F_{n+1} 的 F_U 分量是同相的，仅 F_V 进行倒相，所以 F_{n+1} 应是 F_n 以 U 轴为基准的一个镜像，如图 3.5 所示。

接收机为了按色度信号原来的相位正确重现色调，必须将倒相的 PAL 行色度信号 F_{n+1} 再重新倒回到 F_n 的位置上。

逐行倒相的目的主要是克服 NTSC 制的色度信号的相位敏感性，PAL 制并没有减小相位失真，只是采用逐行倒相的方法使相邻两行相位失真产生两种方向相反的色调畸变，其合成矢量相加补偿，从而减小了相位失真。但其不利因素在于可能导致合成矢量的幅度减小，表现为色饱和度下降。

此外，PAL 电视信号接收机采取两行色度信号电平平均，当相邻两行的色度信号有差别时，平均的结果必然导致两行各自模糊。

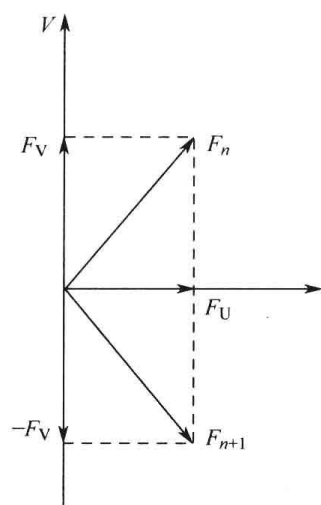


图 3.5 PAL 制色度信号矢量图

3.2 数字电视技术基础

现代视频会议系统其本质上是一种数字电视系统，也可以说现代视频会议系统的技术基础是数字电视技术，充分而翔实地了解数字电视技术，对于我们掌握视频会议的基本原理和评价方法有着积极的促进作用。

图 3.6 是数字电视系统的基本构成图，系统由节目生成、信号处理、信号传输和信号接收四部分组成，本节将重点介绍第二部分，即与视频会议技术相关的原理，包括信号采样、信源压缩编码、系统复用。

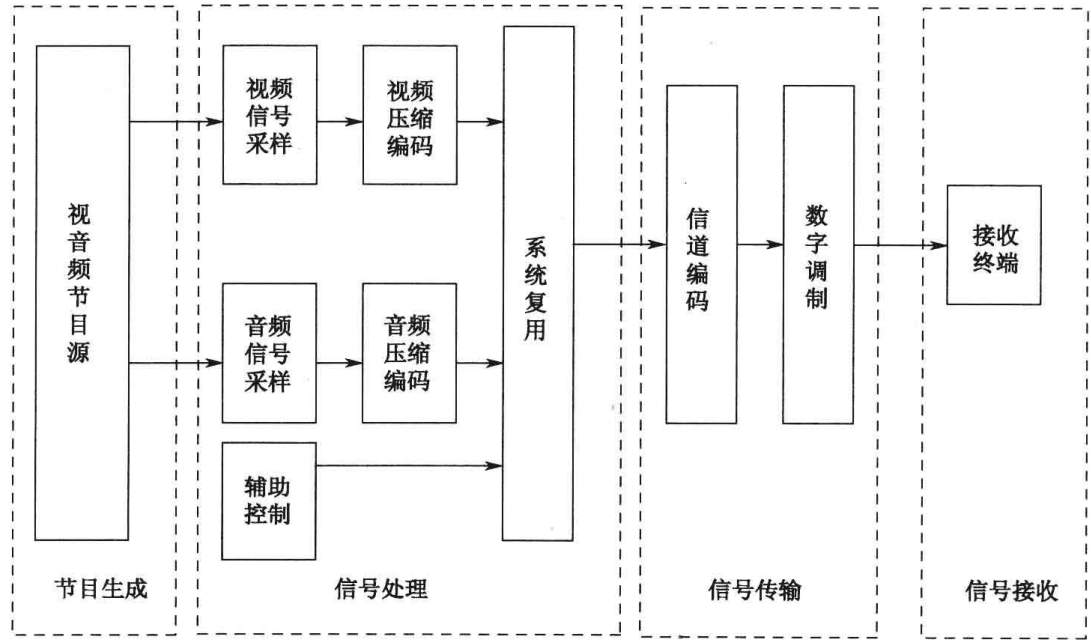


图 3.6 数字电视系统基本构成图

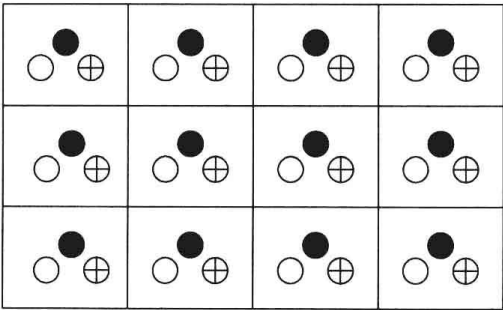
3.2.1 视频信号采样格式

根据电视信号的构成，在对模拟电视信号进行数字采样编码时，有两种相应的方法，方法一是对复合全电视信号进行采样的复合编码；方法二是对亮度信号和两个色差信号进行独立的采样的分量编码。复合编码相对分量编码虽然简单实用，但分量编码消除了不同电视制式的差别，使不同的模拟电视制式在数字电视上得到了统一，为交流电视节目提供了方便，因此分量编码已经成为国际上通用的数字电视信号采样编码方式。

我国标准清晰度电视（SDTV）信号遵循推荐性国家标准 GB/T14857-1993《演播室数字电视编码参数规范》，该规范等效采用 ITU-R BT.601-3 建议书，该规范建议亮度信号 Y 的采样频率为 525/60 和 625/50 两系统行频公倍数 2.25 MHz 的 6 倍，即 13.5 MHz。两个色差信号 U 和 V 的采样频率为亮度信号采样频率的一半，即 6.75 MHz。在这里我们可以看到，亮度信号的采样频率和其他两个色度信号的采样频率之比为 4:2:2，为此我们将其称之为 4:2:2 格式。

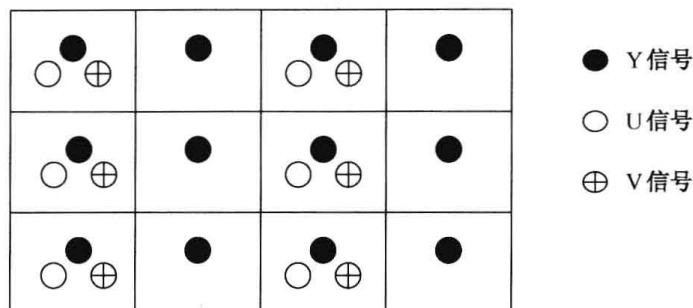
YUV 信号有许多采样格式，常见的是 4:4:4、4:2:2 和 4:2:0 三种。4:4:4 格式是指 Y、U、V 三个信号的采样频率均为 13.5 MHz，该格式一般应用于对图像质量要求比较高的视频设备。4:2:0 格式是指 Y 信号采用 13.5 MHz 的采样频率，U、V 信号采样频率为亮度信号采样频率的一半，即 6.75 MHz。其与 4:2:2 格式的区别在于虽然 U、V 信号均采样，但在实际传送数据时，U、V 信号数值交替传输。

在 YUV4:4:4 格式下，只是 RGB 色彩空间的对应变换，没有数据信息的丢失，是无损的。在 4:2:2 和 4:2:0 格式下牺牲了一部分色差信息，这样做的目的是充分利用了人眼的对亮度敏感而对色彩不敏感的特性，有利于降低传输数据量的大小。YUV 采样格式如图 3.7 所示。

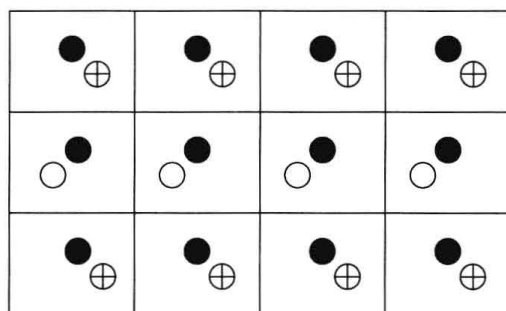


(a) YUV4:4:4 格式

图 3.7 YUV 采样格式比较



(b) YUV4:2:2 格式



(c) YUV4:2:0 格式

图 3.7 YUV 采样格式比较 (续)

3.2.2 视频信号的参数

1. 标清电视和高清电视的图像像素

我国标准清晰度电视系统规定：对亮度分量的采样频率为 13.5 MHz，那么 625 行 50 场制式每行有 $864 \{ (13.5 \times 10^6 \text{ 个采样 / 秒}) / (15 \text{ 625 行 / 秒}) \} = 864$ 亮度采样 / 行。实际显示行为 575 行，取偶数后 576 行，考虑画面宽高比 4:3，则有效采样数 $576 \times \frac{4}{3} = 768$ ，最终规定为图像水平方向的采样点数为 720，图像像素为 720×576。

我国高清晰度电视 (HDTV) 信号遵循广电推荐的行业标准 GY/T155-2000 《高清晰度电视节目制作及交换用视频参数值》，该标准参照 ITU-R BT.709-3 建议书。该规范建议亮度信号 Y 的采样频率为 1125/50/2:1 HDTV 电视系统行频的公倍数 2.25 MHz 的 33 倍，即 74.25 MHz。两个色差信号 U 和 V 的采样频率为亮度信号采样频率的一半，即 37.125 MHz。每行有 $2 \text{ 640} \{ (74.25 \times 10^6 \text{ 个采样 / 秒}) / (28 \text{ 125 行/秒}) \} = 2 \text{ 640}$ 亮度采样/行。高清晰度电视实际显示行 1 080 行，考虑画面宽高比 16:9，则水平方向的分辨力为 $1 \text{ 080} \times \frac{16}{9} = 1 \text{ 920}$ ，图像像素为 1 920×1 080。

2. 数字电视清晰度

数字电视的清晰度用电视线为单位来表示, 通常分为垂直清晰度、水平清晰度和斜向清晰度。

(1) 垂直清晰度

电视图像水平线条数, 在理想的条件下, 与电视有效的扫描行数相同。

$$L_v = N_v \quad (3.10)$$

式中, L_v 为垂直清晰度, N_v 为有效扫描行。

(2) 水平清晰度

电视图像垂直线条数, 为水平方向上, 与图像有效高度相等的宽度内, 所具有的有效像素数。

$$L_H = N_H \times (h/w) \quad (3.11)$$

式中, L_H 为水平清晰度; N_H 为水平有效像素; h 为图像高; w 为图像宽。该水平清晰度是根据采样定理限定的图像水平清晰度值, 但由于实际的视频信号标称的带宽一般低于采样频率的一半, 因此视频信号标称带宽对应的水平清晰度为

$$L'_H = L_H \times \frac{2 \text{倍视频标称带宽}}{\text{采样频率}} \quad (3.12)$$

式中, L_H 为根据采样定理限定的图像水平清晰度值。

(3) 斜向清晰度

$$L_\theta = L_H / \sin \theta \quad (3.13)$$

式中, L_θ 为斜向清晰度, L_H 为水平清晰度, θ 为与水平方向的夹角。

3. 几种常见数字电视信号主要参数

根据上述章节, 表 3.1 给出了几种常见数字电视信号的主要参数。

表 3.1 数字电视信号视频参数

数字电视系统	576i	1080i	480i	720P
类型	标清	高清	标清	高清
隔行比	2:1	2:1	2:1	1:1
场频 / Hz	50	50	59.94	59.94
帧频 / Hz	25	25	29.97	59.94
行频 / kHz	15.625	28.125	15.734	44.955
每帧总行数	625	1125	525	750

续表

图像宽高比	4:3	16:9	4:3	16:9
采样定理理论带宽 / MHz	6.75	37.125	6.75	37.088
信号标称带宽 / MHz	6	30	4.2	30
R、G、B、Y 信号采样频率 / MHz	13.50	74.25	13.50	74.176
色差信号采样频率 / MHz	6.75	37.125	6.75	37.088
R、G、B、Y 信号有效行数 (4:2:2)	576	1 080	480	720
R、G、B、Y 信号每行总采样数 (像素)	864	2 640	858	1 650
R、G、B、Y 信号每行有效采样数 (像素)	720	1 920	720	1 280
有效像素数	720×576	1 920×1 080	720×480	1 280×720
采样定理限定的图像水平清晰度理论值 (电视线)	540	1 080	540	720
视频标称带宽限定的图像水平清晰度理论值 (电视线)	480	872.73	336	582.4
像素宽高比	1.07	1.00	0.889	1.00

数字电视视频信号通常使用下列格式来表示：

每帧有效扫描行数 / 帧频或场频 i (隔行扫描) 或 p (逐行扫描)

例如，1 080/50i 表示有效扫描行 1 080 行，场频 50 Hz，隔行扫描。720/60P 表示有效扫描行 720 行，帧频 60 Hz，逐行扫描。这样我国和欧洲的 PAL 制可表示为 576/50i，美国的 NTSC 制可表示为 480/60i，我国的高清电视体制可表示为 1 080/50i。

4. 视频信号的码率

在完成对电视信号的采样之后，接下来的工作是要对其进行量化编码，目前采用的量化编码位数一般为 8 或 10。

以 576/50i 和 480/60i 标清清晰度电视标准为例：

- 当采用 10 比特量化编码时，亮度信号的码率为取样频率×量化比特数 = $13.5 \text{ (MHz)} \times 10 \text{ (bit)} = 135 \text{ Mbps}$ ；
- 2 个色差信号的码率为 $2 \times 6.75 \text{ (MHz)} \times 10 \text{ (bit)} = 135 \text{ Mbps}$ ，总的码率为亮度信号码率+色差信号码率 = 270 Mbps

以 1080/50i 高清清晰度电视标准为例：

- 当采用 10 比特量化时，亮度信号的码率为取样频率×量化比特数 = $74.25 \text{ (MHz)} \times 10 \text{ (Bit)} = 742.5 \text{ Mbps}$ ；
- 2 个色差信号的码率为 $2 \times 37.125 \text{ (MHz)} \times 10 \text{ (bit)} = 742.5 \text{ Mbps}$ ，总的码率为亮度信号码率+色差信号码率 = $742.5 + 742.5 = 1\,485 \text{ Mbps}$ 。

由上述可知,高清晰度电视的码率是标准清晰度电视的 5.5 倍。

由于电视信号在水平和垂直消隐期间内没有视频信息,因此,一般用有效码率来表示在单位时间内与视频信号有关的数据量,当使用磁带、硬盘或光盘存储数字视频信号时,可以只记录有效码率代表的视频信息。

以 480/60i 标清清晰度电视标准为例:

- 当采样格式 4:2:2, 量化编码为 8 比特时,亮度信号的码率为每行的取样点数 $\times$ 有效扫描行数 $\times$ 量化比特数 $\times$ 帧频 $= 720 \times 480 \times 8 \times 30 = 82.944 \text{ Mbps}$;
- 2 个色差信号的有效码率为 $2 \times 360 \times 480 \times 8 \times 30 = 82.944 \text{ Mbps}$, 总的有效码率为亮度信号有效码率 + 色差信号有效码率 $= 82.944 + 82.944 = 165.888 \text{ Mbps}$ 。

576/50i 标清清晰度电视的有效码率也为 165.888 Mbps, 两者相同。

以 1 080/60i 高清晰度电视标准为例:

- 当采样格式 4:2:2, 量化编码为 8 比特时,亮度信号的码率为每行的取样点数 $\times$ 有效扫描行数 $\times$ 量化比特数 $\times$ 帧频 $= 1\,920 \times 1\,080 \times 8 \times 30 = 497.664 \text{ Mbps}$;
- 2 个色差信号的有效码率为 $2 \times 960 \times 480 \times 8 \times 30 = 497.664 \text{ Mbps}$, 总的有效码率为亮度信号有效码率 + 色差信号有效码率 $= 82.944 + 82.944 = 995.328 \text{ Mbps}$;
- 1 080/50i 的有效码率为 829.44 Mbps

可见,高清晰度电视 1 080/60i 的有效码率是标准清晰度电视 480/60i (NTSC 制)的 6 倍,高清晰度电视 1 080/50i 的有效码率是 576/50i (PAL 制)的 5 倍。

现有的传输手段根本不可能传输如此高的码率,因此必须将采样编码后的数字信号进行大幅度的压缩。

3.2.3 视频压缩编码技术

1. 冗余类型

视频图像中含有自然的冗余度,一般而言,冗余类型主要有空间冗余、时间冗余、信息熵冗余三种。

(1) 空间冗余

空间冗余是静态图像中存在的最主要的数据冗余,例如,一幅图像中有较大的背景区域,在此区域中所有点的光强、色彩,以及饱和度都是十分相近的,这种空间连贯性就成为空间相关或空间冗余。

(2) 时间冗余

时间冗余是视频序列中经常包含的冗余，序列图像中的相邻帧往往包含相同或类似的背景和运动物体，只不过运动物体所在空间位置略有不同。这种相邻帧间数据的高度相关性就称为时间冗余。

(3) 信息熵冗余

信息熵冗余也称编码冗余，为表示图像数据的一个像素点，只要按其信息熵的大小分配相应的比特数即可。对于实际图像数据的每个像素，在图像获取时很难得到它的信息熵，一般是对每个像素采用相同的比特数来表示，这样就必然存在冗余，这种冗余称为信息熵冗余。

2. 视频编码器采用的主要技术

为了去除上述的这三种主要冗余，视频编码器包括很多编码算法，这些算法在编码器中被有效地组合在一起，使整个编码器具有较高的压缩效率。目前主流的视频编码器采用的技术主要有预测、变换、量化、扫描和熵编码，这些技术在编码器中的基本次序如图 3.8 所示。

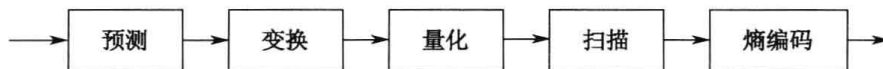


图 3.8 视频压缩关键技术流程图

(1) 预测

预测编码即是利用信号间的相关性，用前面一个或多个信号作为当前信号的预测值，对当前信号的实际值与预测值的差进行编码，经过预测后，残差信号的能量比源信号的能量低，从而降低了待编码信号的信息量。预测编码是消除视频冗余的重要编码工具之一，针对空间冗余有帧内预测编码，对于时域冗余有基于运动估计技术的帧间预测编码。

帧内预测利用图像在空间上相邻像素之间具有相关性的特点，由相邻像素预测当前块的像素值，可以有效去除块间冗余。

帧间预测是消除运动图像时间冗余的技术。运动图像邻近帧中的场景存在一定的相关性，因此，可为当前块搜索出在邻近参考帧中最相似的预测块，并根据预测块的位置，得出两者之间的空间位置的相对偏移量，即通常所说的运动矢量。根据运动矢量，从指定的参考帧中找到预测块的过程称为运动补偿。对于运动矢量的编码通常采用与周围块的运动矢量进行差分预测来编码。通过搜索得到的运动矢量的过程称为运动估计。运动估计是编码器中帧间预测的核心模块，根

据运动估计处理的基本单位进行划分, 运动估计有基于像素的运动估计、基于块的运动估计、基于网格的运动估计和全局运动估计等形式。

帧间预测技术在预测方向上从最简单的前向预测又发展出了后向预测和双向预测, 即 I、P、B 三种帧类型, 即 I 帧 (Intra Picture)、P 帧 (Predicted Picture) 和 B 帧 (Bidirectional Picture), 如图 3.9 所示。

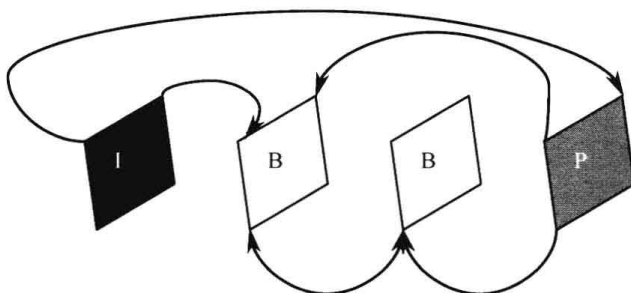


图 3.9 帧间预测模式

在预测精度上, 从整像素预测到 $1/2$ 像素预测以及 $1/4$ 像素预测。甚至 $1/8$ 、 $1/16$ 像素预测; 预测模式采用从 16×16 块到 8×8 块、 4×4 块等多级划分。

(2) 变换

变换被认为是视频编码中最有效的技术之一, 将空域信号变换到频域信号, 有效地去除了信号的相关性, 并使大部分能力集中到低频区域。常用的变换技术有很多种, 如 K-L 变换、傅里叶变换、离散余弦变换 (DCT)、小波变换。下以常用的 DCT 变换为例加以说明。

DCT 将每个图像块表示为二维余弦函数的加权和, 通过这些余弦变换函数与图像块的像素的加权和操作, 可以将像素映射到频域上, 得到的频域系数能量更集中、更易于压缩。

例如, 图 3.10 (a) 是一个 8×8 图像块, 包含了 64 个像素值, 这些像素值具有很强的相关性, 经过 8×8 DCT 变换, 得到 64 个变换系数如图 3.10 (b) 所示, 其中最左上角的变换系数为 DC (直流) 系数, 其余的变换系数为 AC (交流) 系数。

得到的 64 个变换系数具有很明显的能量分布特征, DC 系数的值非常显著, 而靠近 DC 系数的低频 AC 系数值比较显著, 距离 DC 系数较远的高频区域的 AC 系数值则不显著, 大部分接近于零。

在视频编码中, 根据人眼视觉特性对高频信息不敏感的特点, 通过量化操作使大部分高频系数量化为零, 最后得到一个 8×8 稀疏变换量化系数块, 这样编码时就大大减少了编码的比特数。

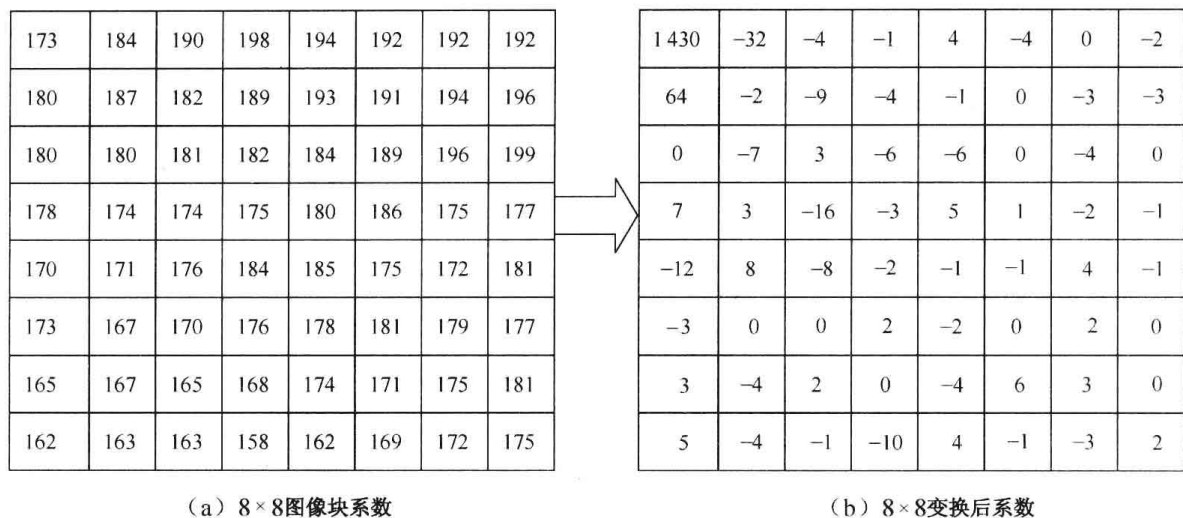


图 3.10 DCT 编码特性示意图

(3) 量化

变换后的量化是降低数据表示精度的过程，通过量化可以减少需要编码的数据量，达到压缩数据的目的。量化减小非 0 系数的幅度以及增加 0 值系数的数目，在一定的主观保真前提下，丢掉那些对视觉效果影响不大的信息，量化是图像质量下降的最主要原因。

量化是在 8×8 像素块上完成 DCT 变换之后进行的，一旦非重要的分量被去除，是无法恢复的，因此量化过程是不可逆的有损压缩过程。量化后的比例系数中数值较大的被映射到非零整数，数值较小的系数被映射到零。

(4) 扫描

变换量化系数在熵编码之前通常要通过“Z”字（Zigzag）扫描。通过 Zigzag 扫描可将二维变换量化系数重新组织为一维系数序列，经过重排序的一维系数再经过有效的组织能够被高效地编码。扫描的顺序一般根据待编码系数的非零系数分布，按照空间位置出现非零系数的概率从大到小排序。排序的结果是使非零系数尽可能地出现在整个一维系数序列前面，而后面的系数尽可能为零或接近于零，这样排序非常利于提高系数的熵编码效率。图 3.11 给出了 8×8 变换量化系数块的 Zigzag 扫描的顺序。

从 DCT 二维系数生成的 64×1 一维数组中的第一个元素代表 DC 系数，其余的 63 个系数代表 AC 系数。这两类系数完全不同，足以将它们分开。

(5) 熵编码

熵编码的目标是去除信源符号在信息表达上的表示冗余。视频编码常用的熵编码方法有三种：变长编码（Variable Length Coding, VLC）、算术编码（Arithmetic

Coding, AC) 和行程编码。变长编码的基本思想是, 为出现概率大的符号分配短码字, 为出现概率小的符号分配长码字, 从而达到总体平均码字最短。代表性的变长编码是哈夫曼 (Huffman) 编码。

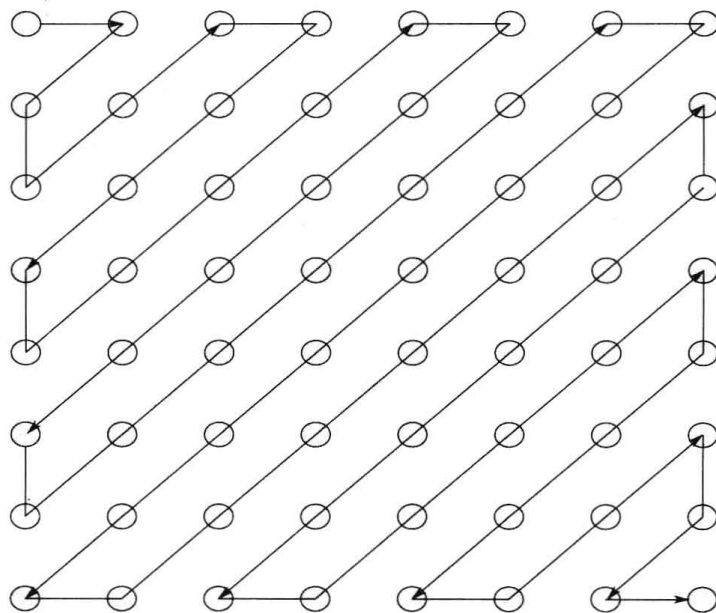


图 3.11 8×8 块 Zigzag 扫描顺序

算术编码和变长编码不同, 不是采用一个码字代表输入信息符号的方法, 而是采用一个浮点数来代替一串输入符号。算术编码计算输入符号序列的联合概率, 将输入符号序列映射为实数轴上的一个小区间, 区间的宽度等于该序列的概率值, 之后在此区间内选择一个有效的二进制小数作为整个符号序列的编码码字。可以看到算术编码是对输入符号序列进行操作, 而非单个符号, 因此平均意义上可以为单个符号分配小于 1 的码字。

行程编码的主要思想是在信源符号和码字之间建立明确的一一对应关系, 以便在恢复时可以准确滴再现原信号, 行程编码本质上是一种无损编码, 常用 RLE (Run-Length Encoding) 表示。

行程编码一般包含两项, 第一项用一个符号代替具有相同值的连续符号, 第二项用来记录原始数据中有多少个这样的值。

如 22233333344444440000000099999 的行程编码可以表示为 (2,3)(3,6)(4,7)(0,8)(9,5), 这样使编码长度小于自然编码的长度。

3.2.4 视频压缩编码标准

视频编码技术经历了 60 年的发展, 在这一过程中, 逐渐形成了变换编码、预测编码、熵编码三类经典技术, 分别用于去除视频信号的空间冗余、时间冗余以

及信息熵冗余。基于这三类技术,逐渐形成了以块为单位的预测+变换的混合编码框架。到目前为止,出现的所有视频编码标准均是基于这一技术框架,包括 ITU-T 的 H.261、H.263、H.264 编码标准,ISO/IEC 的 MPEG-1、MPEG-2、MPEG-4 视频编码标准以及中国的 AVS 编码标准。

ISO/IEC 下的 MPEG (Motion Picture Experts Group) 专家组专门负责制定多媒体领域内的相关标准,主要应用于存储、广播电视、因特网或无线网上的流媒体。MPEG 组织在 1992 年制定了面向 VCD 应用的 MPEG-1 标准,速率为 1.5 Mbps。1994 年发布了面向 DVD、数字视频广播等应用的 MPEG-2 标准,适用于 1.5~60 Mbps 甚至更高码率。1998 年制定了面向低码率的 MPEG-4 标准。

ITU-T 则主要制定面向实时视频通信领域的视频编码标准,如视频电话、视频会议等应用。开始于 1984 年的 H.261 是 MPEG-1 标准的前驱,主要视为在 ISDN 上实现可视电话、视频会议标准而制定的。在 H.261 的基础上,1996 年 ITU-T 制定了 H.263 标准,相继又推出了 H.263+和 H.263++等。

2001 年 ITU-T 和 MPEG 联合成立了 JVT (Joint Video Team) 工作组,制定了一个新的视频编码标准,在 ISO/IEC 称为 MPEG-4 的第 10 部分 (MPEG-4 Part 10 AVC),在 ITU-T 中称为 H.264 标准。

中国于 2002 年 6 月成立的 AVS (Audio Video Coding Standard) 工作组负责为国内多媒体工业界制定相应的数字音视频编码国家标准。

下面简要介绍几类主要的视频编码标准。

1. H.261/263 标准

H.261 标准是第一个被产业化的视频编码标准,它主要针对双向视频会议应用。H.261 算法复杂度较低,便于可视电话进行实时编解码,同时由于双向视频传输对延迟较敏感,因此 H.261 仅允许采用帧内预测图像 (I 帧) 与前向预测图像 (P 帧),而不允许使用双向预测图像 (B 帧)。H.261 采用基于 8×8 块的离散余弦变换进行残差信号的变换编码。DCT 把图像的每个 8×8 像素块变换到频域,产生 64 个频率成分 (第一个系数称为 DC 系数,其他的称为 AC 系数)。为了量化 DCT 系数, H.261 对所有的 AC 系数采用固定的线性量化,对量化的系数进行行程编码。

H.263 在 H.261 之后得到开发,主要是以更低的比特率实现更高的编码质量。其基本原理继承了 H.261,但提供了许多新的编码选项,如 4 个 8×8 亮度块运动矢量、重叠块运动补偿、去块效应滤波器、多参考帧预测等。由于 H.263 算法在一般情况下优于 H.261 的编码效率,它成为电视会议首选的编码算法。后来,ITU-T 又发布了 H.263+和 H.263++更新的版本,提供了更多的编码工具选项来提高编码效率,同时增强了容错抗误码的能力。

2. MPEG-1/MPEG-2/MPEG-4 标准

MPEG-1 标准是 MPEG 工作组开发的第一个视频编码标准, 主要应用在数字媒体动态图像的存储和检索, 如采用 CIF 分辨率的 VCD。MPEG-1 与 H.261 相似, 不过需要编码器具有更高的性能, 以便支持视频内容的较高运动型而不是典型的可视电话功能。与 H.261 不同, 在 MPEG-1 中允许采用 B 帧。MPEG-1 支持逐行视频, 不支持隔行视频。

MPEG-2 标准是专门针对数字电视而开发的, MPEG-2 既能满足逐行视频的需求, 又能够满足电视领域常用的隔行视频的要求。MPEG-2 标准是由 MPEG-1 发展而来的, 它的运动矢量在水平和垂直两个方向上允许时半像素精度, 参考图像以插值方式得到更高的分辨率。这种方法提高了运动补偿的精度, 使帧间预测更为准确, 显著提高了运动估计的性能。MPEG-2 在码率达到 4~8 Mbps 时即可满足消费类视频应用的需求, 如数字卫星电视、数字有线电视、DVD 及高清电视。

MPEG-4 延续 MPEG-2 的成功, 提供基于对象的编码功能, 注重用户与内容的交互性, 实现各种新工具以支持编码对象即视频之间的融合。但由于 MPEG-4 基于对象的编码功能过于复杂, 并未在产品中受到重视, 相关实现主要集中在改善低比特率压缩效率以及提高容错性能上。MPEG-4 SP (Simple Profile) 以 H.263 为基础, 为改善压缩效率增加了一些新的编码工具, 例如, 不受限运动矢量、自适应帧内 DC/AC 预测等。另外, MPEG-4 标准增加了大量容错功能, 以支持丢包情况下的恢复。随后推出的 MPEG-4 ASP (Advanced Simple Profile) 以 MPEG-4 SP 为基础, 增加了与 MPEG-2 类似的 B 帧及隔行工具, 以及 1/4 像素运动补偿和全局运动补偿的选项。

3. H.264/AVC 标准

在过去几年, 视频编码技术最重要的发展之一是 ITU-T 和 ISO/IEC 联合视频工作组 JVT 开发的 H.264/AVC 标准, 该标准在 ISO/IEC 中的正式名称是 MPEG-4 Part 10/AVC, 在 ITU-T 中的正式名称是 H.264, 因此被常被称为 H.264/AVC 标准。

和之前的编码标准相比, H.264/AVC 的编码效率有较大幅度的提高。在相同重建图像质量下, H.264/AVC 比 MPEG-2 编码效率高 2~3 倍, 比 H.263 和 MPEG-4 高 1.5~2 倍。H.264/AVC 性能的提升是以计算复杂度的增加为代价的, 其编码的计算复杂度大约相当于 H.263 的 3 倍, 解码复杂度相当于 H.263 的 2 倍。

H.264/AVC 和以前的标准一样, 也是采用混合编码架构, 但其引入了新的编码工具, 提高了编码效率。新的编码工具包括帧内空域预测、整数变换、多参考帧预测、可变块大小预测、1/4 精度像素运动补偿等技术。

为了更好地支持网络传输，H.264/AVC 引入了面向 IP 包的编码机制，将编码流分成视频编码层（Video Coding Layer，VCL）和网络提取层（Network Abstraction Layer，NAL）两个层次。VCL 保存视频压缩后的数据流，NAL 主要是为 VCL 提供一个与网络无关的统一接口，采用统一的数据格式对视频数据进行封装打包后使其在网络中传输。

H.264/AVC 标准包括 Baseline Profile (BP)、Extended Profile (EP)、Main Profile(MP)、High Profile(HP)四个层次，其压缩编码效率逐步提高，但相应的复杂度和对硬件的性能指标要求也相应提高，如图 3.12 所示。

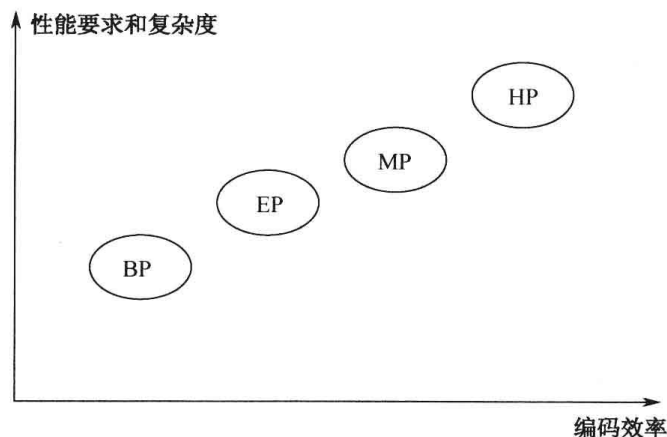


图 3.12 H.264 层次能力

在 H.264 发展初期，业界主要关注 H.264/AVC 的 BP 和 MP。BP 对计算能力以及系统内存的需求较低，并且针对低时延进行了优化，它非常适合可视电话、移动电视以及其他低成本、低延时的应用。MP 采用了 B 帧编码以及复杂的熵编码器，虽然可以提供更高的压缩效率，但其要求的处理能力比 BP 高很多。MP 主要用于广播域内容存储等应用，以满足更高压缩效率和视频质量的需求。为了应对电影以及工作室级高保真度编码的需求，JVT 又推出了以 HP 为代表的一系列高端应用。一些针对性优化工具被引入，如量化矩阵、色度空间编码等，进一步提升了 H.264/AVC 的压缩性能。

H.264/AVC 因其具有更高的压缩比、更好的 IP 和无线网络信道的适应性，在数字视频通信和存储领域得到越来越广泛的应用，如有线电视、卫星电视、视频会议和远程视频监控等。

4. AVS 标准

AVS（Audio Video Coding Standard）是由我国独立制定的具有自主知识产权的视频编码标准，它由中国音视频标准工作组组织制定。AVS 工作组成立于 2002 年，其目的是要建立中国自己的而数字音 / 视频国家标准。和以往的视频编码标

准相比, AVS 编码效率有了突破性的提高, 其与 H.264/AVC 标准有着类似的结构, 和国际上的 H.264/AVC 标准编码性能相当, 但实现复杂度明显降低。可以被利用在高分辨率数字广播、高密度激光数字存储媒体、无线宽带多媒体通信、互联网宽带流媒体等重要的信息产业领域。

3.2.5 系统复用

在模拟电视系统中, 多路模拟信号是用频分复用 (FDM) 方式复合在一起的。FDM 方式的缺点是各信号间易相互干扰, 而且频谱利用率比较低。

而在数字电视系统中, 系统层编码将一个或多个视频流、音频流、数据流、节目管理信息以及同步信息以时分复用 (TDM) 方式复合成一个串行码流, 进行存储或传输, 其系统复用的过程如图 3.13 所示。

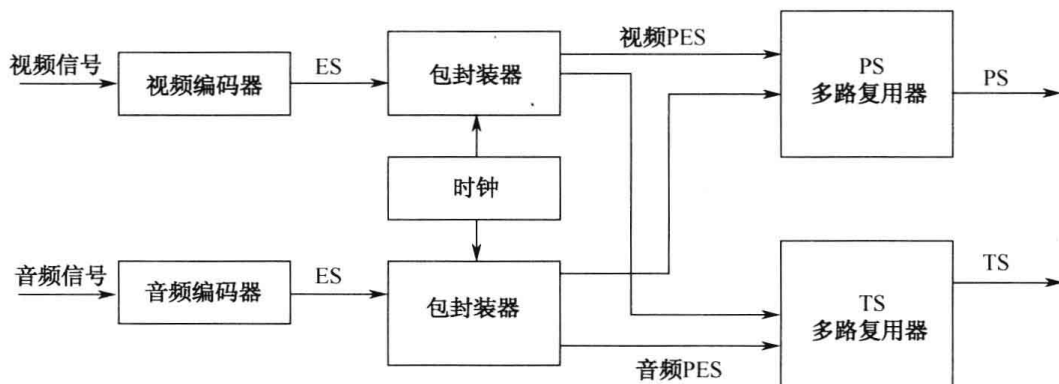


图 3.13 数字电视流系统复用

① 一路节目的视、音频及其他辅助数据经数字化后, 通过压缩层完成信源压缩编码, 分别形成基本码流 ES。

② 系统层将不同的基本流 ES 分别加包头打包为 PES 包, PES 又称为分组基本码流。这一步骤在打包器内实现, PES 的长度可在一定范围内变化。

③ PES 再与编码的辅助数据流复用成系统流。对于光盘等不易产生误码的信道, PES 与辅助数据流复用成节目流 (PS), PS 码率可变或固定。

对于数字卫星电视等易产生误码的信道, PES 和辅助数据及控制数据复用成传送流 (TS)。TS 由一系列的 TS 包组成。为了在包失步后便于重新进入包同步状态, 包长固定为 188 Byte, 由包头 (4 Byte) 和需要传送的有用信息即负载 (184 Byte) 组成, 如图 3.14 所示。

每个 TS 包由包识别符 (PID) 标识, 从 PID 可以判断其后面净荷的数据类型, 是视频、音频流、节目专用信息 (PSI) 还是其他数据包。TS 的码率通常固定不变, 自适应调整字段或放置节目参考时钟 (PCR), PCR 以固定频率插入包头, 表示编码端的时钟, 解码端根据 PCR 调整时钟, 保证正确解码。

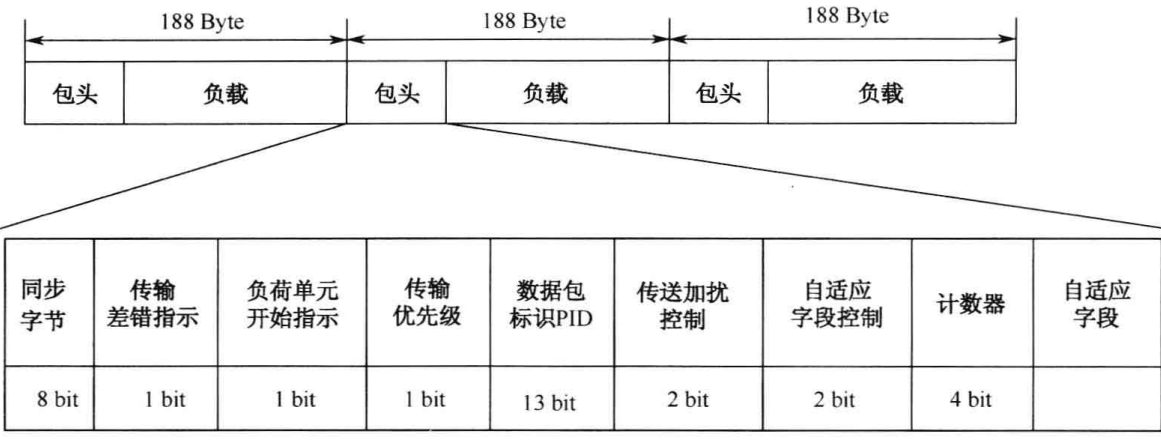


图 3.14 TS 码流语法结构

3.3 主要视频接口特性

本节介绍广播电视领域目前主要使用的视频接口，这些接口也是视频会议设备所广泛应用的设备接口，包括 CVBS 接口、S-Video 接口、YPbPr/YCbCr 色差接口、DVI 接口、HD-SDI 接口。

1. 复合视频接口（CVBS）

复合视频（Composite Video Broadcast Signal）接口用于视频设备之间的标清模拟视频信号的连接。复合视频接口也称 AV 接口或 Video 接口，是一种普遍使用的模拟视频接口。

它是一种音、视频分离的视频接口，结构类型为圆形单触点同轴连接器，又称 RCA 插头，标识为黄色。视频信号没有经过其他信号处理过程，信号保真度较好，图像品质受使用的线材影响较大。当用于数字显示设备时，需要一个模数转换过程，将损失一定信噪比，所以一般数字显示设备不建议使用。

2. S-Video 接口

S-Video 接口是一种很常见的视频接口，S 代表 Separate，与复合视频接口相比，它将亮度和色度分开传输，不再进行亮度和色度信号的混合传输，从而避免了混合视频信号输出时亮度信号和色度信号之间的相互干扰而产生的图像失真，极大地提高了图像的清晰度。但 S-Video 接口仍要将两路色差信号混合为一路色度信号，进行传输后再将其分离为色差信号进行处理，这样多少会产生一定的失真，而且由于色差信号的混合导致色度信号的带宽也有一定的限制。

3. YPbPr/YCbCr 色差接口

YPbPr/YCbCr 色差接口是在 S-Video 接口基础上,把色度信号中的色差信号分开传输,其中 YPbPr 表示逐行扫描色差输出,YCbCr 表示隔行扫描色差输出。色差输出将 S-Video 接口中的色度信号分解为色差 Cr 和 Cb,这样就避免了两路色差信号混合译码并再次分离的过程,也保持了色度信道的最大带宽,只需要经过反矩阵译码电路就可以还原为 RGB 三原色信号而成像,这就最大限度地缩短了视频源到显示器成像之间的视频信号信道,避免了因烦琐的传输过程所带来的影像失真,色差接口是目前最好的模拟视频接口。支持 480i、576i、576p、720p、1 080i 等标清和高清视频格式。

4. DVI 接口

DVI 的全称为 Digital Visual Interface,即数字视频接口。DVI 接口与 VGA 接口都是计算机中最常用的接口,由于计算机内部传输的是二进制的数字信号,如果使用 VGA 接口连接液晶显示器,就需要先把信号通过显卡中的 D/A 转换器转变为 R、G、B 三原色信号和行、场同步信号,这些信号通过模拟信号线传输到液晶显示器内部还需要相应的 A/D 转换器将模拟信号再一次转变成数字信号,才能在液晶显示器上显示出图像来。在上述的 D/A、A/D 转换和信号传输过程中不可避免会出现信号的损失和受到干扰,导致图像出现失真甚至显示错误,而 DVI 传输的是数字信号,数字图像信息不需经过任何转换,就会直接被传送到显示设备上,避免了信号的损失,减少了数字→模拟→数字烦琐的转换过程,大大节省了时间,因此它的速度更快,有效消除拖影现象,使图像的清晰度和细节表现力都得到了大大提高。

DVI 接口有多种规范,常见的是 DVI-D (Digital) 和 DVI-I (Intergrated)。DVI-D 只能传输数字信号,可以用它来连接显卡和数字平板电视。DVI-I 则不仅支持数字信号,还可以支持模拟信号。

5. HD-SDI 接口

SDI (Serial Digital Interface) 接口是一种“数字分量串行接口”,它把数据字的各个比特通过单一通道顺序传送。由于串行数字信号的数据率很高,在传送前必须经过处理,对原始数据流进行扰频,并变换为不归零倒置 (NRZI) 码,以确保在接收端可靠地恢复原始数据。这样在概念上可以将数字串行接口理解为一种基带信号调制。

SDI 接口不能直接传送压缩后数字信号,数字录像机、硬盘等设备记录的压缩

信号重放后，必须经解压并经 SDI 接口输出才能进入 SDI 系统。如果反复解压和压缩，必将引起图像质量下降和延时增加。

HD-SDI 接口是一种广播级的高清数字输入和输出端口，其中 HD 表示高清信号。在非编后期制作，在广电领域等领域，HD-SDI 应用较为广泛。

SDI 对应的速率为 270 Mbps，适合以 576/50i 和 480/60i 标清清晰度电视的传输。HD-SDI 对应速率为 1.485 Gbps，适合 1080/50i 高清清晰度电视的传输。

6. HDMI 接口

HDMI 接口的全称是 High-Definition Multimedia Interface，中文意思为高清多媒体接口。HDMI 接口可以提供高达 5 Gbps 的数据传输带宽，可以传送没有经过压缩处理的音频信号及高清视频信号，同时无须在信号传送前进行数 / 模或者模 / 数转换，可以保证最高质量的影音信号传送。

HDMI 接口和 DVI 接口采用了相同的传输机制，所以 HDMI 对 DVI 拥有非常强大的兼容性。因为只涉及接口的形状、尺寸和引脚定义，在电路部分没有任何的变化，因此可通过 HDMI-DVI 转换进行两种接头间的物理转换。HDMI 标准中也考虑了和 DVI 的设备兼容问题，只要 HDMI 设备检测到对方发送的信号中不包含 HDMI 标准中规定的特殊控制数据，就会把对方确认为 DVI 设备，并把传输规格切换到 DVI 格式，从而保证了良好的兼容性。

此外，具备 HDMI 接口的设备具有“即插即用”的特点，信号源和显示设备可以自动协商合适的视 / 音频格式。

第4章 视频会议网络技术

现代视频会议网络技术是建立在 ITU-T 的 H.32X 系列标准基础之上的，比较常见的有 H.320、H.323、H.324 标准。

H.320 是基于电路交换网络的视频会议网络技术标准，也称为专线方式的视频会议系统。H.320 规定了视频和语音编 / 解码标准、复用与控制等一整套协议。H.320 语音编 / 解码标准有 G.711、G.722、G.728、G.729，视频编 / 解码标准有 H.261、H.263、H.263+H.264，复用部分采用 H.221、H.230，控制部分采用 H.242。

H.323 由 H.320 发展而来，是 ITU 在 1996 年制定的，称为“不保证服务质量的局域网上的多媒体视听系统标准”。该协议提供了在没有传输质量保证的网络上进行多媒体通信的规范，主要针对 IP 网络环境。H.323 中语音编解码标准有 G.723.1、G.711、G.722、G.728、G.729，视频编 / 解码标准有 H.261、H.263、H.263+H.264，复用部分采用 H.225、H.230，TCP/IP，控制部分采用 H.245，适用于 LAN 和 Internet。

H.324 是 1996 年颁布的基于 PSTN 的视频会议系统，因为传输速率比较低（33.3 kbps），视频质量比较差，因而没有很大发展，而后期扩展的针对无线网络环境的要求，称为 H.324M。

除上述三个标准之外，还有 H.321、H.322 标准，H.321 标准是将 H.320 适配到 ATM 环境的技术标准，而 H.322 是建立在有服务质量保证的局域网上的多媒体通信标准。H.32X 各类标准对比如表 4.1 所示。

表 4.1 H.32X 系列标准框架

协议标准	H.320	H.321	H.322	H.323	H.324
传输网络	窄带电路交换网络	ATM	有 QoS 保障的分组交换网络	无 QoS 保障的分组交换网络	PSTN 或无线环境
视频标准	H.261/263/264	H.261/263/264	H.261/263/264	H.261/263/264	H.261/263/264
音频标准	G.711/722/728	G.711/722/728	G.711/722/728	G.711/722/723 /728/729	G.723
复用协议	H.221	H.221	H.221	H.225.0	H.223
控制协议	H.230/242	H.242	H.230/242	H.245	H.245

由于 H.323 是在基于 IP 的网络环境中实现可靠的面向音视频实时应用的，经过多年的发展和完善，H.323 已经被广泛地接受和应用，称为事实上的主流视频会

议网络技术标准，目前大规模的商用视频会议系统均使用 H.323 技术标准。

此外，IETF 的 SIP 协议也是视频会议网络的重要标准之一，单纯的 SIP 系统与 H.323 相比，在系统组成和呼叫流程方面有很多相似之处。首先，SIP 的 UA 等价于 H.323 的终端，实现呼叫的发起和接收，并完成所传输媒体的编解码应用；SIP 代理服务器、重定向服务器以及注册服务器的功能等价于 H.323 的网守，实现终端的注册、呼叫地址的解析以及路由。其次，虽然在呼叫信令和控制的具体实现上不同，但一个基于 SIP 的呼叫流程与 H.323 的 Q.931 类似，SIP 采用的会话描述协议（SDP）则类似于 H.323 的呼叫控制协议 H.245。

SIP 的 Internet 背景，使得 SIP 在构造基于端到端的应用和复杂的拓扑结构时较 H.323 要灵活很多。表 4.2 为 H.323 与 SIP 协议对比。

表 4.2 H.323 与 SIP 协议对比

协议类型	H.323	SIP
应用目标	基于 IP，实现面向可靠的音视频和数据的实际应用	基于 Internet 的实时音视频应用
组织结构	由网守、MCU 等组成，以 H.245 为呼叫控制协议	SIP 代理服务器、重定向服务器、注册服务器构成，以 SDP 为会话控制协议
实现难易度	基于二进制编码，系统自由度较小，如要实现大规模应用，则需要对整个网络的各个环节进行规划	基于文本编码，系统自由度较大，对于构成大规模应用具有优势

重要的是，随着移动通信技术和各类宽带接入技术的发展，迫切要求电信运营商推出与接入无关的多网融合的视频会议业务，而采用 SIP 协议的 IMS 网络技术恰好顺应了这种要求，这对发展统一融合的视频会议业务具有积极的现实意义。

基于上述观点，本章着重从以下方面来进行论述。

- 采用 H.323 协议的 IP 视频会议网络技术；
- 采用 SIP 协议的 IMS 视频会议网络技术；
- 基于 IMS 技术的视频会议。

4.1 H.323 视频会议网络技术

H.323 协议标准规定了在无服务质量（QoS）确保的分组网络（Packet-Based Network, PBN）上的多媒体通信系统的技术要求，用于在分组交换网络中提供多媒体业务。H.323 协议标准定义了 H.323 系统的组成部分，包括终端（Terminal）、用于地址翻译和访问控制的网守（Gatekeeper）、多点会议控制单元（MCU）、多点会议控制器 MC、多点处理器（MP），以及介于电路交换网和分组交换网之间

的 H.323 网关 (Gateway) 等实体。这些组件组合在一起, 提供点对点或多点间多媒体会议的通信服务能力。H.323 系统的网络组成结构如图 4.1 所示。

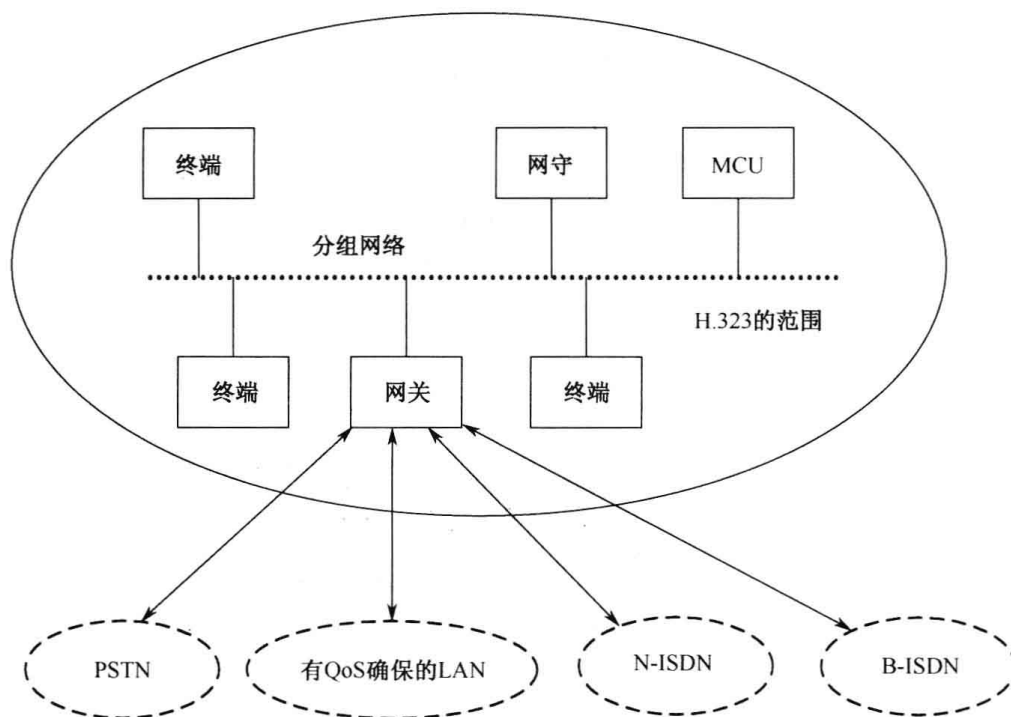


图 4.1 H.323 系统的网络组成结构

(1) 终端

在 H.323 系统中, H.323 终端提供了点对点或多点会议中音视频和数据的通信能力, 它在接入网守的控制下完成呼叫的建立与释放, 接收对端发送来的视 / 音频编码信号, 解码后播放给用户, 并将近端的原始多媒体信号编码后经由视讯会议业务网进行交换, 因此它在多媒体通信系统中担任关键的角色。H.323 终端可以是一台 PC 或者一个独立的嵌入式操作系统设备。图 4.2 描述了 H.323 终端设备的功能框架图。

从图 4.2 中可以看出, H.323 终端包含了多个功能单元, 如视频编解码器、音频编解码器、用户数据应用、H.225.0 层、系统控制功能以及分组网络的接口。下面逐一介绍各单元功能。

- 视频编解码器: 对来自视频源 (如摄像机) 的视频进行编码发送, 对接收到的视频码进行解码并输出到视频显示器。
- 音频编解码器: 对来自麦克风的音频信号进行编码发送, 对接收到的音频码进行解码并输出到扬声器。
- 用户数据应用: 支持远程信息处理应用, 如电子白板、静态图像传输、文件交换、数据库访问、音频图形会议等。

- 系统控制单元 (H.245 和 H.225.0): 为 H.323 终端的正确操作提供信令, 提供呼叫控制、能力交换、命令和指示的信令及消息。
- H.225.0 层 (H.225.0): 对发送的视频、音频、数据和控制流进行格式化, 形成消息输出到网络接口; 从网络接口接收到的消息中提取视频、音频、数据和控制流。另外, 对每一种媒体类型, 完成适当的逻辑成帧、顺序编号、差错检测和差错纠正。
- 网络接口: 因不同实现而异, 超出 H.323 协议标准的范围。然而, 网络接口必须提供在 H.225.0 建议中描述的服务, 包括 H.245 控制信道、数据信道和呼叫信令信道服务; 必须提供可靠的 (如 TCP) 端到端服务。对音频信道、视频信道和 RAS 信道, 必须提供非可靠的 (如 UDP) 端到端服务。服务可以是双工或单工、单播或多播, 取决于应用、终端能力和网络配置。

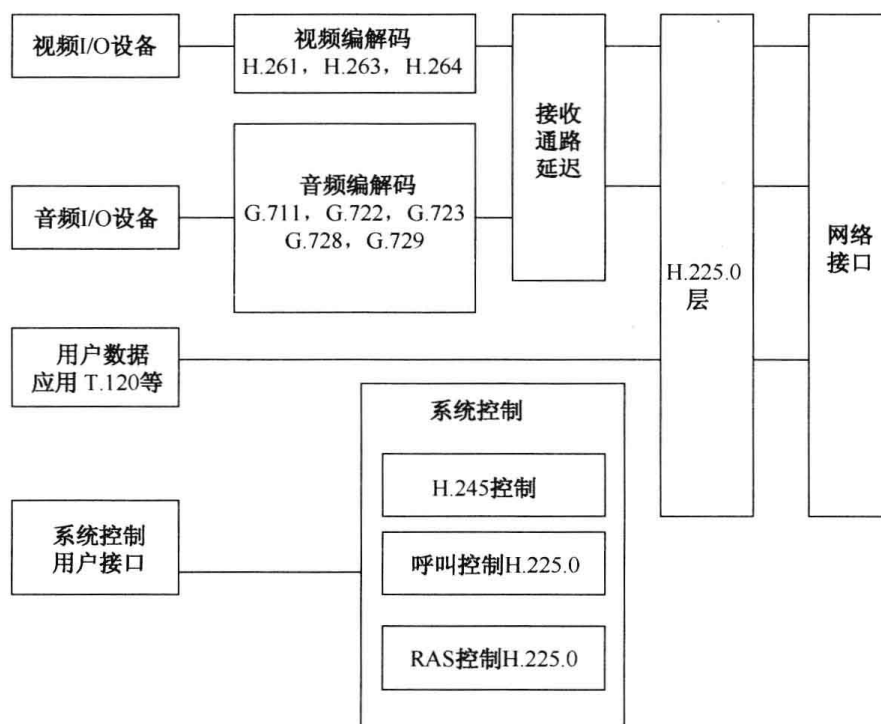


图 4.2 H.323 终端的功能框架图

(2) 网守

在 H.323 系统中, 网守主要为 H.323 端点提供呼叫控制服务, 是 RAS 协议的实现实体。网守的基本功能如下。

- 地址转换: 通过到达网守的登记消息建立翻译表, 根据此表提供别名地址翻译成传输层地址的服务。由于网守随时可能接收新的登记消息, 所以翻译表是不断更新变化的。
- 接入许可控制: 根据 H.323 端点的访问权限和使用带宽大小等情况, 利用 H.225.0 标准的许可请求、许可确认和许可拒绝 (ARQ/ACF/ARJ) 消息来

发送、接收、确认或拒绝呼叫。

- 带宽控制：利用 H.225.0 标准的带宽请求、带宽拒绝和带宽确认（BRQ/BRJ/BCF）消息完成对端点占用带宽资源的动态分配。
- 域管理：对登记在域内的端点提供服务。
- 可选的服务：包括呼叫控制、呼叫批准、呼叫控制信令、带宽管理等。

虽然网守在逻辑上和其他端点分离，但也可以将网守的功能融入 H.323 终端、网关和多点控制单元等物理设备中。由单一网守管理的所有终端、网关和多点控制单元的集合称之为 H.323 域。

（3）多点控制单元（MCU）

多点控制单元支持三个或以上节点设备的会议，在 H.323 系统中，一个多点控制单元由一个多点控制器（MC）和几个多点处理器（MP）组成，但可以不包含 MP。

MC 提供针对多点会议的控制功能，主要包括：

- 通过与每个端点进行能力交换，确定会议的能力集，在有端点加入或离开会议时，修改能力集。
- 根据能力集，决定会议中各点的选定通信模式。
- 选择集中、分散或混合的会议方式。

MP 从端点接收视频、音频和数据流，处理后返送回端点。视频处理包括视频交换和视频混合，音频处理包括音频交换、音频混合以及两种结合处理。

（4）网关

网关是 H.323 系统的一个可选组件。在 H.323 会议节点终端与其他 ITU 标准相兼容的终端通信时，由网关完成它们之间的互通。网关的主要功能包括提供 H.323 网络与非 H.323 网络之间的信令消息格式和内容转换、通信协议流程转换和媒体流格式转换等。

4.1.1 H.323 协议栈

H.323 协议栈如图 4.3 所示，其中低三层为分组网协议，第四层传输层包含可靠传送（TCP）和不可靠传送（UDP）两类协议，之上均为应用层，分为声像应用、数据应用、终端控制和管理三类。

H.323 协议栈利用了许多已有的成熟标准，定义了自身协议框架，具体来说，由以下几个部分组成。

- 语音编码采用相应的 G 系列建议，其中 G.711（PCM）为必备的编码方式，其余为任选方式，包括 G.722、G.728、G.729 和 G.723.1 等。

- 视频编码采用 H.260 系列建议，如 H.261、H.263、H.264 等。
- 音频和视频编码信号均封装在 RTP（Real time Transport Protocol，实时传输协议）协议分组中，提供定时信息和数据报序号，供接收端重组信号。RTCP（Real time Transport Control Protocol，实时传输控制协议）协议是 RTP 协议的一部分，提供 QoS 监视功能。

应用层					
声像应用	终端控制和管理				数据应用
G.7xx H.26x	RTCP	H.225.0 终端至网守 信令 (RAS)	H.225.0 呼叫信令	H.245 媒体信道控制	T.120 系列
RTP					
不可靠传送协议 (UDP)			可靠传送协议 (TCP)		
网络层					
链路层					
物理层					

图 4.3 H.323 协议栈

- 数据通信采用 T.120 系列建议，是用于多媒体会议的数据协议。
- RAS（Registration Admission and Status）信令协议是 H.225.0 协议的一种，是端点（终端、网关和 MCU）和网守之间使用的协议，其主要作用是为网守提供确定的端点地址、状态、执行呼叫接纳控制等功能。
- H.225.0 呼叫信令协议是 H.323 的核心协议，主要用于呼叫控制。在任何呼叫开始之前，首先必须在端点之间建立呼叫联系，同时建立 H.245 控制信道。
- H.245 是一种通用的多媒体通信控制协议，主要针对会议通信设计。H.323 采用 H.245 作为控制信令协议，用于控制媒体信道的建立、维护和释放。

需要重点讨论的协议是 RAS 信令协议、H.225.0 呼叫信令协议和 H.245 控制信令协议，这三个协议定义了 H.323 实体之间相互交换的消息类型和格式。H.323 协议栈主要引用这三个协议。

4.1.2 RAS 信令协议

1. RAS 消息

RAS 信令协议是 H.225.0 协议的一种，它是网守用来控制其所在域内端点

(包括终端、网关、MCU)的信令协议。RAS 信令是通过发送不同的消息来完成的,主要执行管理功能,它包含以下几个过程。

(1) 网守搜索 (Gatekeeper Discovery)

网守搜索用于端点搜索其所属的网守,采用多播机制,其后所有 RAS 消息均限定在端点和其归属网守之间传送。

(2) 端点登记 (Endpoint Registration)

端点登记用于端点向网守登记其自身信息,主要是别名和呼叫控制信道传输层地址,包括去除登记的过程。

(3) 端点定位 (Endpoint Location)

端点定位用于端点或网关向相应的网守询问某一端点呼叫控制信道的传输层地址。

(4) 呼叫接纳 (Call Admission)

呼叫接纳是发起呼叫的第一步操作,询问网守是否允许该呼叫发起。

(5) 呼叫退出 (Call Disengagement)

呼叫结束后通知网守,该端点已经退出呼叫(恢复空闲)。

(6) 带宽管理 (Bandwidth Management)

带宽管理支持端点在呼叫过程中提出带宽改变的要求,由网守作决定。

(7) 状态查询 (Status)

状态查询主要用于网守询问终端的开机/关机状态。

(8) 网关资源指示

网关资源指示向网守通告网关的可用资源。

在不含网守的网络环境中,H.225.0 呼叫信号信道直接建立在两个通信的端点之间;而在含有网守的网络环境中,H.225.0 呼叫信道由网守来决定呼叫建立是在网守和端点之间还是直接建立在端点之间。

RAS 信令消息列表如表 4.3 所示。

表 4.3 RAS 信令消息

过程	消息	英文全称	消息中文名称	消息类型
网守搜索	GRQ	Gatekeeper Request	网守搜寻请求	请求
	GCF	Gatekeeper Confirm	网守搜寻证实	响应
	GRJ	Gatekeeper Reject	网守搜寻拒绝	响应

续表

过程	消息	英文全称	消息中文名称	消息类型
端点登记	RRQ	Registration Request	登记请求	请求
	RCF	Registration Confirm	登记证实	响应
	RRJ	Registration Reject	登记拒绝	响应
	URQ	Unregistration Request	注销请求	请求
	UCF	Unregistration Confirm	注销证实	响应
	URJ	Unregistration Reject	注销拒绝	响应
端点定位	LRQ	Location Request	端点定位请求	请求
	LCF	Location Confirm	端点定位证实	响应
	LRJ	Location Reject	端点定位拒绝	响应
呼叫接纳	ARQ	Admission Request	呼叫接纳请求	请求
	ACF	Admission Confirm	呼叫接纳证实	响应
	ARJ	Admission Reject	呼叫接纳拒绝	响应
呼叫退出	DRQ	Disengage Request	呼叫退出请求	请求
	DCF	Disengage Confirm	呼叫退出证实	响应
	DRJ	Disengage Reject	呼叫退出拒绝	响应
状态查询	IRQ	Info Request	信息请求	请求
	IRR	Info Request Response	信息响应	响应
	IACK	Info Acknowledgement	对 IRR 消息的证实	响应
	INAK	Information Negative Acknowledgement	对 IRR 消息的拒绝	响应
带宽管理	BRQ	Bandwidth Request	带宽请求	请求
	BCF	Bandwidth Confirm	带宽证实	响应
	BRJ	Bandwidth Reject	带宽拒绝	响应
网关资源指示	RAI	Resource Availability Indication	网关资源可用性指示	请求
	RAC	Resource Availability Confirmation	网关资源可用性证实	响应

2. RAS 消息流程

(1) 网守搜索

网守搜索是端点用来决定在哪一个网守上注册的过程。网守搜索有两种方式：静态方式和自动方式。静态方式通过端点配置完成，将其归属网守的传输层地址预先置入配置文件或初始化文件中。自动方式允许端点和归属网守的关系可以随时间而改变，当原有网守出故障时可以自动切换到其他网守上。在自动方式中，端点采用搜索多播地址发送 GRQ 消息，询问“谁是我的网守？”，可以有一个或多个网守回送 GCF 消息，表示“我可以是你的网守”，并在消息中告诉端点它的 RAS 地址，拒绝该端点在其上登记的网守则返回 GRJ 消息。如果多个网守回

送 GCF 消息,端点可在其中任选一个作为其归属的网守。如果超时仍未收到网守的响应,端点可重发 GRQ,两次相邻的 GRQ 之间时间间隔不小于 5 s。如果仍未收到响应,则改用静态配置。如果端点发现所属网守登记失效,即网守发送拒绝登记请求或不回送任何响应,就必须重新启动搜索过程。网守搜索的过程为如下所述。

- 端点在启动后,首先向网守发送 GRQ 消息,寻找网守。该端点发送的 GRQ 消息中包含端点类型、端点自身的 RAS 地址、希望在其上登记的网守标识等参数。
 - 网守对端点信息进行分析,确定是本区域端点,发 GCF 确认。
 - 如果网守不愿意该端点在其上登记,则返回 GRJ 消息,并给出拒绝原因。
- 网守搜索如图 4.4 所示。

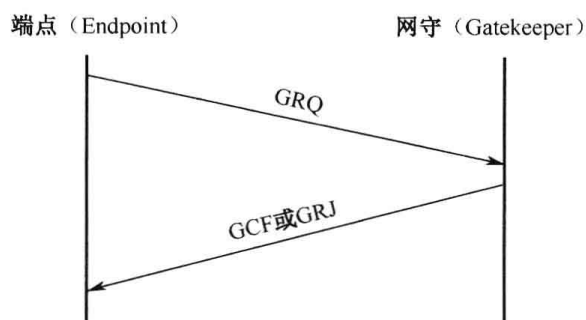


图 4.4 网守搜索

(2) 端点登记与注销

端点登记须根据网守的 RAS 地址向其发送 RRQ 消息,消息包含的最重要的两个参数就是端点别名及其信令传输层地址,别名可以为 E.164 地址或 H.323 标识。E.164 地址可由接入码和电话号码组成,该接入码可以用来标识网关。H.323 标识为字符串形式,可以是用户名、E-mail 或其他标识名。一个端点可以有多个别名,所有别名都应在 RRQ 消息中送往网守,它们将被翻译为同一个信令传输层地址。一般情况下,网守将返回 RCF 消息,告之网守的信令传输层地址,并将端点别名和信令传输层地址登记到翻译表中。消息中有一个“生存时间”参数,单位为秒,指示本次登记信息的寿命。如果网守收到 RRQ,其别名和以前 RRQ 消息相同,但传输层地址不同,则回送 RRJ 消息,并指明拒绝原因是“重复登记”。如果传输层地址和以前相同,但别名不同,则作为更新登记消息修改其翻译表。如果端点在 RRQ 中没有包含别名,网守可以赋予一个别名,然后将此别名由 RCF 回送给终端。如果端点想改变其别名和地址的对应关系,可以先将以前的登记信息由 URQ 消息删除,网守回送 UCF 消息;如果网守发现该端点原来并未在它上面登记,则回送 URJ 消息。网守也可以向端点发送 URQ 消息,表示要删除该端点的登记消息,端点须回送 UCF 消息,其后,端点须重新登记才能发起呼叫,一般

要在新的网守上登记。

端点登记与注销流程如图 4.5 所示。

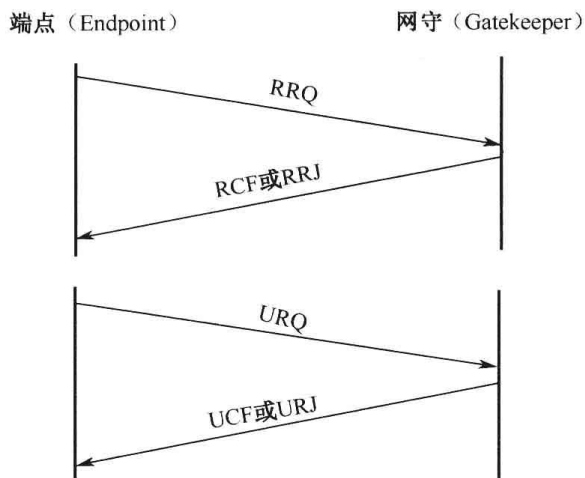


图 4.5 端点登记与注销流程

(3) 呼叫接纳与退出

ARQ/ACF 和 DRQ/DCF 是整个呼叫控制过程第一对和最后一对消息，分别标志呼叫的开始和结束。在 ARQ 中，端点给出目的地信息（E.164 地址或 H.323 标识）以及所要求的带宽。网守如果同意接纳此呼叫，则回送 ACF，其中包括的两项主要参数是允许分配的带宽和翻译后所得的目的地呼叫信令传输层地址或者是网守本身的呼叫信令传输层地址。不但主叫发呼叫时需要请求网守接纳，被叫收到呼叫建立信令时也要向网守发送 ARQ 请求，因此在 ARQ 中有一个“应答呼叫”（Answer Call）参数，指示是否来话呼叫，具体流程如下所述。

- 当端点发起呼叫时，端点向网守发送 ARQ 请求用户接入认证 / 地址解析。在 ARQ 消息中，端点给出目的地信息以及所要求的带宽。
- 网守如果同意接纳此呼叫，则回送 ACF 消息，其中包括的两项主要参数是允许分配的带宽和翻译后所得的目的地的呼叫信令传送层地址或者是网守本身的呼叫信令传输层地址。
- 网守如果不同意接纳此呼叫，则回送 ARJ 消息拒绝。
- 呼叫完毕，端点向网守发送 DRQ 请求退出呼叫。
- 一般情况下，网守回送 DCF 进行确认。
- 如果网守不同意该端点退出呼叫，则回送 DRJ 消息拒绝。

呼叫接纳与退出如图 4.6 所示。

从上述介绍的消息流程可以得知，RAS 信令协议的另一个主要功能是提供呼叫连接双方端点的呼叫信令地址，即用于传输呼叫信令消息的网络传输地址。在 IP 地址环境中，网络传输地址为 IP 地址+端口号。由于有了网守，呼叫双方可以用别

名 (E.164 地址或 H.323 标识符)，而不是用呼叫双方的呼叫信令地址进行呼叫。

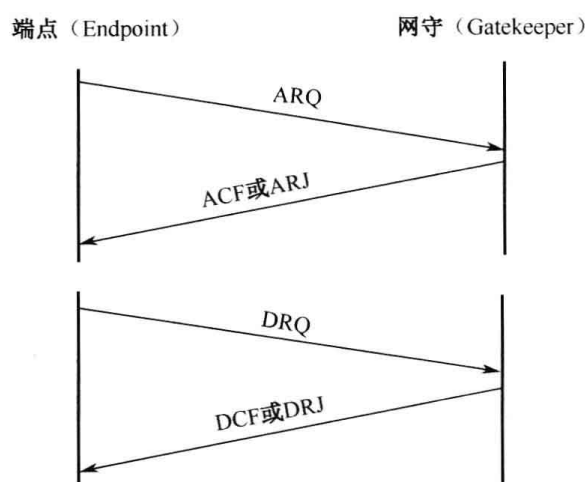


图 4.6 呼叫接纳与退出

4.1.3 H.225.0 呼叫信令协议

H.225.0 呼叫信令协议用于建立 H.323 端点设备（包括终端、网关、MCU）之间的连接，信令消息通过 H.225.0 呼叫信令通道传输，呼叫信令的作用是在端点之间建立和拆除呼叫，呼叫信令消息在可靠的 TCP 通信信道上进行传递。

1. H.225.0 呼叫信令消息

H.225.0 呼叫信令消息是按 H.225.0 建议修改过的 Q.931/Q.932 消息，Q.931/Q.932 消息是 ITU-T 制定的一种关于呼叫控制的标准，是 ISDN 用户网络接口第三层关于基本呼叫控制的描述。

H.225.0 借鉴了 Q.931/Q.932 的长处，并简单地对它的消息进行重用，同时由于 H.225.0 呼叫信令消息不承担连接控制任务，为了适应 H.323 协议的总体体系结构，对 Q.931/Q.932 消息进行了精简。

H.225.0 呼叫信令消息如表 4.4 所示。

表 4.4 H.225.0 呼叫信令消息

消息类型	消息名
呼叫建立消息	Setup
	Setup Acknowledge
	Call Proceeding
	Alerting
	Connect
	Progress

续表

消息类型	消息名
呼叫清除消息	Release Complete
其他消息	Information
	Notify
	Status Enquiry
	Status
	Facility
	User Information (用户信息)

(1) 初始化 (Setup)

如果一个端点想与另一个端点建立呼叫，则它发送的第一个呼叫信令消息就是初始化消息。如果该端点使用网守提供的服务，那么只有在收到网守发送的 ACF 消息表示网守允许它建立呼叫后，它才能够发送初始化消息。在 Setup 消息中含有呼叫类型、呼叫信令地址等，还可以包含主叫当前的通信能力。

(2) 激活 (Altering)

这一消息由被叫端点发出，表示被叫用户处于激活状态。被叫端点告诉主叫端点已经把呼入电话通知了被叫方的用户，即被叫正在振铃。

(3) 呼叫在处理中 (Call Proceeding)

这条消息由被叫端点发出。被叫端点收到初始化消息就发送这一消息，表示请求的呼叫正在被处理，并且不能接受新的呼叫。Call Proceeding 信令中最重要的信息是 H.245 地址，供呼叫双方建立 H.245 信令信道时使用。

(4) 进行中 (Progress)

这一消息由被叫端点发送，用来表明呼叫正在进行。在与电路交换网络的互操作中，这一消息尤显的重要。其中的“原因”信息元素尽管是可选的，必要时可以用作传递网内语音或通知的补充。

(5) 连接 (Connect)

连接消息由被叫实体发往主叫实体，其作用是表明被叫方已经接受了主叫方发起的呼叫。当被叫的用户拿起听筒后，被叫端点发出此消息告诉对方，呼叫已经建立完毕。至此，呼叫信令消息的功能已经基本完成，应该由 H.245 消息进行下一步的控制。

(6) 完全释放 (Release Complete)

此消息被用来结束一个呼叫。呼叫双方都可以发出这条消息来表示拆除呼叫

信令连接。拆除呼叫信令信道并不表示结束呼叫，因为呼叫双方进行实际通信的信道是由 H.245 控制信道建立的，只要这些信道还存在，呼叫就可以继续。如果呼叫信令信道和各个 H.245 信道都关闭了，则通信结束。

(7) 性能 (Facility)

在 H.225.0 基本呼叫控制信令中，性能 (Facility) 消息用来指示呼叫重定向，包括呼叫前转、呼叫转向、强制将呼叫控制转向网守三种情况。当主叫收到此消息时，意味着被叫不愿意接受该呼叫，而是让主叫去呼叫别的端点。主叫则只能用 Release Complete 消息拆除呼叫信令连接，然后根据 Facility 消息的内容向别的端点发出新的呼叫。

2. H.225.0 呼叫消息流程

下面以两端点在同一网守上登记为例，分别介绍直接呼叫和网守选路路由呼叫两种方式的信令过程，以便更加深入地了解 H.323 系统的呼叫建立流程。

(1) 直接呼叫信令过程

其信令过程如图 4.7 所示，简要说明如下。

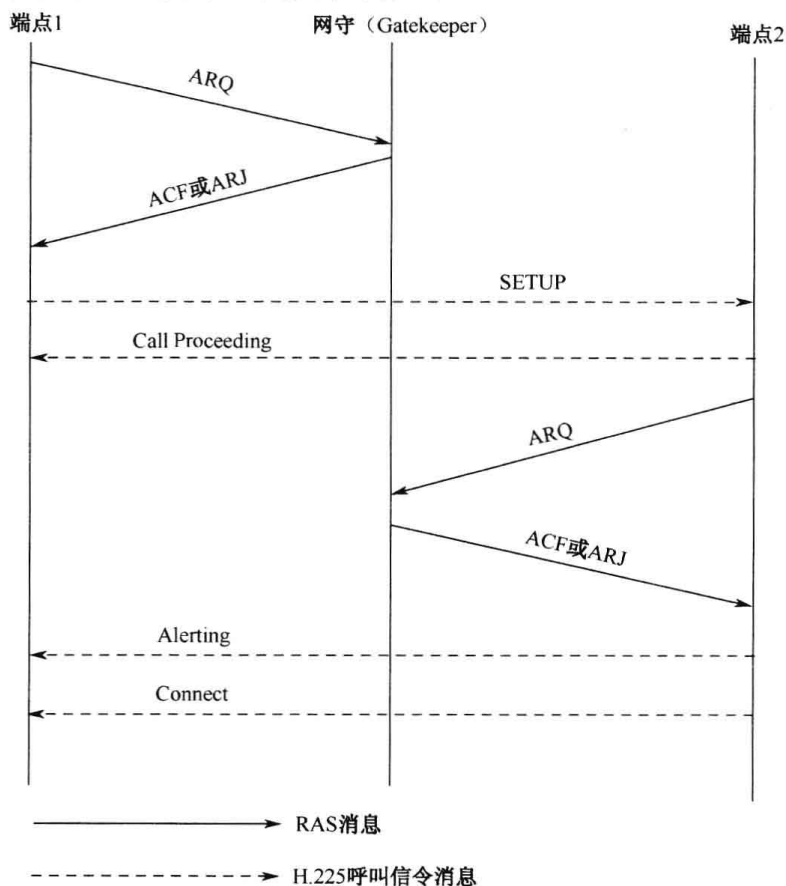


图 4.7 两个端点在同一网守上注册——直接呼叫信令过程

① 端点 1（主叫）在 RAS 信道上向其网守发送 ARQ 消息，请求发起至端点 2 的呼叫。

② 网守同意接纳此呼叫并翻译出端点 2 的呼叫信令信道传输层地址（IP 地址+TCP 端口号），由 ACF 消息回送端点 1。

③ 端点 1 建立至端点 2 的呼叫信令信道，在此信道上发送 Setup 消息。如果 ARQ 中已带呼叫参考值 CRV，则 Setup 及其后信令消息中的 CRV 应取此相同值，这个呼叫参考值被用于对应同一呼叫的两个实体之间的所有呼叫信令消息中。

④ 端点 2 回送 Call Proceeding 消息，指示呼叫已抵达，正在处理之中。对于两个 H.323 终端之间的呼叫，除了用户信息单元（User-User Information Element, UUIE）之外，消息一般不必带其他信息单元。如果是 H.323 终端和网关之间的通信，则网关将把来自电路交换网侧收到的信息单元，如承载能力和进展表示语透明回传给 H.323 终端。

⑤ 端点 2 经 RAS 信道向网守发送 ARQ，请求接受此呼叫。如果网守同意接纳，回送 ACF。

⑥ 端点 2 向端点 1 回送 Alerting 消息，等待用户应答。

⑦ 用户应答，端点 2 向端点 1 发送 Connect 消息，消息中带有端点 2 的 H.245 控制信道 TCP 端口号。至此，呼叫建立完成。

⑧ 如果网守不同意端点 2 接受此呼叫，则回送 ARJ，此时端点 2 将向端点 1 发送 Release Complete 消息。

（2）网守路由呼叫信令过程

网守路由呼叫信令过程如图 4.8 所示，与直接呼叫信令的差别如下所述。

① 网守向端点 1 回送的 ACF 消息中包含的不是端点 2 的呼叫信令信道传输层地址，而是网守自身的呼叫信令信道传输层地址。同时，网守建立至端点 2 的呼叫信令信道。

② 其后，端点 1 的呼叫信令消息只能发送到网守，再由网守将其转发给端点 2。

由于端点 2 只和网守建有信令信道，因此其信令消息也只能发往网守，再由网守转发给端点 1。

③ 当呼叫建立成功时，端点 2 仍经 Connect 消息告知其 H.245 控制信道传输层地址，但网守向端点 1 发送的 Connect 消息所含信息取决于 H.245 控制消息的传送方式。如果网守决定采用直接方式传送媒体控制消息，则消息中包含的是端点 2 的 H.245 控制信道地址；如果采用网守转接方式，则消息中包含的是网守的 H.245 控制信道地址。

④ 主、被叫任何一端挂机，送 Release Complete 消息给网守，网守再送

Release Complete 消息给对端。主、被叫间断开 TCP 连接。

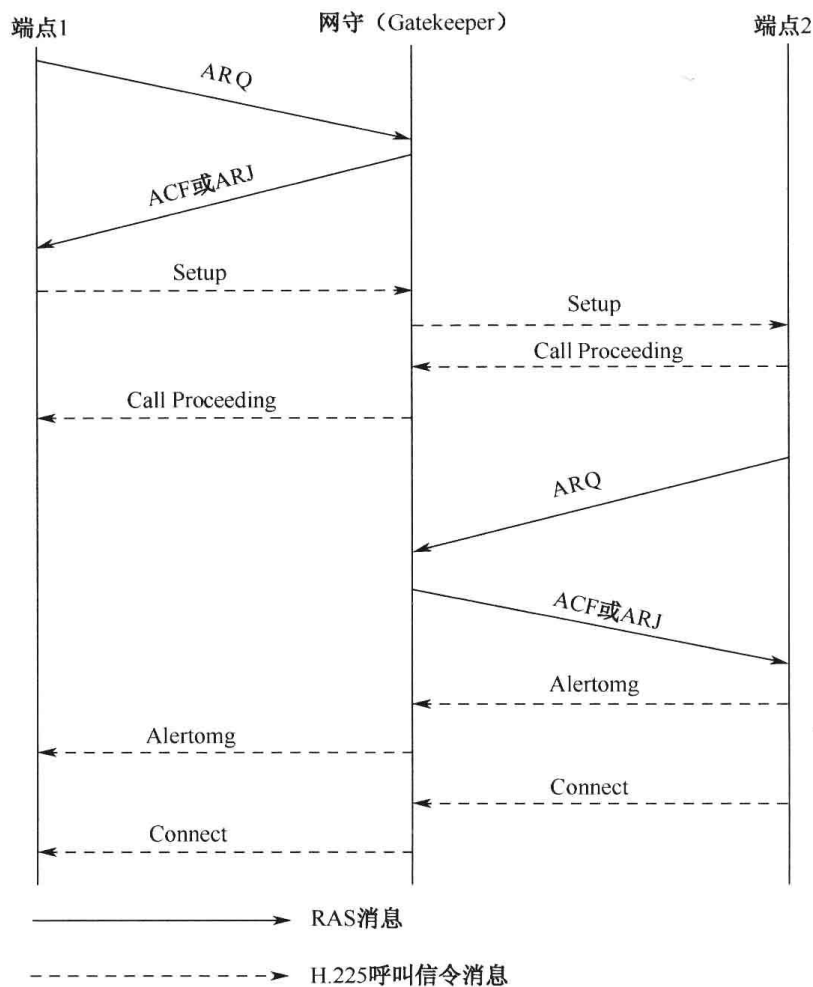


图 4.8 两个端点在同一关守上注册——网守路由选路信令过程

4.1.4 H.245 控制信令协议

H.245 是 H.323 协议族中用于逻辑信道控制的协议，主要完成多媒体通信中每个逻辑通道的建立、维护和释放，具体包括逻辑通道的打开和关闭、参数设定、双方能力协商等控制功能。同时，它还要完成多点会议呼叫中逻辑信道的配合控制功能。

在 H.245 中定义了两类信道：控制信道和通信信道。控制信道必须在任何逻辑信道之前先行建立，并在通信结束后释放。

控制信道也称为 H.245 信道，位于不同 H.323 实体上的两个 H.245 对等信令实体通过该信道传送 H.245 消息，以控制通信信道的建立和释放，控制信道是基于 TCP 连接的可靠信道，连接端口号动态分配。在 H.225.0 呼叫建立过程中，主、被叫端点或 GK 通过 Setup 和 Connect 消息互相交换各自分配的 H.245 端口地址，呼叫建立完成后，H.245 控制信道就建立起来了。每个呼叫有且只有一个

H.245 控制信道，它在整个呼叫期间始终存在，直到呼叫完成后才予释放。

通信信道在 H.245 中称为逻辑信道，在其上传送用户通信信息。一般来说，两个实体间可有多条逻辑信道，在呼叫中可以根据需要随时建立和释放，在 H.245 协议中称为打开和关闭。逻辑信道的开关由 H.245 控制协议完成，每个逻辑信道在打开时赋予一个标识号。可以认为控制信道是一个特殊的永久逻辑信道，其信道号指定为“0”。

H.323 中大多数逻辑信道都是单向信道，在会议通信中尤为如此。但是 T.120 数据通信协议和普通的点到点电话通信都要求双向信道，它由一对单向逻辑信道构成，占用两个逻辑信道号。H.245 逻辑信道打开过程既支持单向信道的建立，也支持双向信道的建立。传送音频和视频信号的逻辑信道为不可靠信道（如 UDP），传送数据信号的为可靠信道（如 TCP），它们的端口号均是动态分配的。所谓逻辑信道的建立，就是通信双方通过逻辑信道打开消息（OpenLogicalChannel）和证实消息（OpenLogicalChannelAck）交换各自分配的端口号。

每个逻辑信道限定传送某种媒体信息、编码算法和数据带宽，因此，在信道建立之前，收发双方必须就这些参数进行协商，确定双方可接收的参数范围，这就是 H.245 的能力交换过程。H.245 协议采用接收方控制的原则建立逻辑信道，发送方只能在接收方规定的范围内确定信道特性参数。能力交换过程的主要功能就是通过适当的消息向对方通告本端的接收能力。该消息也允许通告发送能力，但其作用只是表示本端的一种选择意向，希望对方在确定其接收能力时予以考虑。当获知对方的接收能力后，本端就在此范围内确定其发送方式，启动逻辑信道打开过程。

1. H.245 协议的消息

H.245 消息可分为请求（Request）、响应（Response）、命令（Command）、指示（Indication）4 种类型。请求和响应消息由协议实体使用，构成协议过程。请求消息要求接收方执行所要求的动作，并立即返回响应。响应消息是对请求消息的回复，可以是证实、拒绝或返回请求的结果。命令消息也要求接收方执行指定的动作，但不要求其回送响应。指示消息只是传送信息，不要求接收方执行动作，也不要求其回复响应，通常是指示终端的状态信息。H.245 协议的消息如下所述。

（1）能力集交换

消息名称	消息说明	消息类型
Terminal Capability Set	终端能力集	请求
Terminal Capability Set Acknowledge	终端能力集证实	响应
Terminal Capability Set Reject	终端能力集拒绝	响应
Terminal Capability Set Release	终端能力集释放	指示

(2) 逻辑信道信令过程

消息名称	消息说明	消息类型
Open Logical Channel	打开逻辑信道	请求
Open Logical Channel Acknowledge	打开逻辑信道证实	响应
Open Logical Channel Reject	打开逻辑信道拒绝	响应
Open Logical Channel Confirm	打开逻辑信道确认	指示
Close Logical Channel	关闭逻辑信道	请求
Close Logical Channel Acknowledge	关闭逻辑信道证实	响应

(3) 接收方关闭逻辑信道请求

消息名称	消息说明	消息类型
Request Channel Close	请求信道关闭	请求
Request Channel Close Acknowledge	请求信道关闭证实	响应
Request Channel Close Reject	请求信道关闭拒绝	响应
Request Channel Close Release	请求信道关闭释放	指示

(4) 主从决定

在建立 H.245 信道过程中,可以使用主从决定,也可以不使用,对于 IP 电话, H.245 体制建议不采用此流程。

消息名称	消息说明	消息类型
Master Slave Determination	主从确定请求	请求
Master Slave Determination Acknowledge	主从确定证实	响应
Master Slave Determination Reject	主从确定拒绝	响应
Master Slave Determination Release	主从确定释放	指示

(5) 往返时延确定

消息名称	消息说明	消息类型
Round Trip Delay Request	往返时延请求	请求
Round Trip Delay Response	往返时延响应	响应

(6) 维护环路

消息名称	消息说明	消息类型
Maintenance Loop Request	维护环路请求	请求
Maintenance Loop Acknowledge	维护环路证实	响应
Maintenance Loop Reject	维护环路拒绝	响应
Maintenance Loop Command off	维护环路命令关闭	命令

(7) H.245 基本命令消息

消息名称	消息说明	消息类型
Flow Control	流量控制	命令
Send Terminal Capability Set	发送终端能力集	命令
Encryption	加密	命令
End Session	结束会话	命令
Miscellaneous Commands	其他命令	命令

“发送终端能力集”命令的作用是在发生中断或其他不确定问题的情况下，请求远端终端发送能力集消息告知其发送和接收能力。除非有非常必要的原因，该命令不能重复发送。“加密”命令用来交换加密能力，命令远端在信道上发送初始矢量。“结束会话”表示呼叫结束，终止 H.245 消息传送，通常是呼叫控制过程的最后一个消息。“其他命令”指的是相当于 H.320 中规定的控制可视电话的一些命令。

(8) H.245 基本指示消息

消息名称	消息说明	消息类型
Function Not Understood	功能无法理解	指示
Jitter Indication	抖动指示	指示
H.225.0 Maximum Skew Indication	H.225.0 最大偏斜指示	指示
User Input	用户指示	指示
Miscellaneous Indication	其他指示	指示

其中，“功能无法理解”返回不理解或不支持的请求、响应或命令，不理解的原因可能是对方的 H.245 版本较高或消息中含有非标准的数据单位。“抖动指示”用于向发送端告知接收端计算的逻辑信道抖动量，有助于对端选择视频信道的比特率和缓冲器控制策略，确定定时信息的传送速率。“H.225.0 最大斜偏指示”用于向远端指示指定的两个逻辑信道之间的平均时间偏差，单位为 ms，常用于可视电话中指示视频和音频之间的同步偏差，偏差的原因可包括取样时间、编码器时延和发送缓冲器的时延。“用户指示”消息用于传送 DTMF 信号，即 0~9、*和 #，供和电路交换网互通时使用。“其他指示”指的是用于可视电话的指示信号。

(9) 会议通信相关的 H.245 消息

消息名称	消息说明	消息类型
Conference Request	会议请求	请求
Conference Response	会议响应	响应
Conference Command	会议命令	命令
Communication Mode Request	通信方式请求	请求
Communication Mode Response	通信方式响应	响应
Communication Command	通信方式命令	命令

消息名称	消息说明	消息类型
MC Location Indication	MC 位置指示	指示
Miscellaneous Conference Indication	其他会议指示	指示

其中，会议类消息主要用来控制会议相关操作，如请求参会终端清单、终端标识、会议标识、成为会议主席、退出会议等。会议命令中的退出会议命令用来结束会议，此命令发出后，会议包含的所有呼叫均释放。通信方式类消息主要供 MC 指示各终端各类媒体信道的类型、通信方式（单播或多播）及通信地址。“MC 位置指示”用于主 MC 向其他端点发送其自身地址，以便实施对会议通信的控制。“其他会议指示”主要通告接收终端或其他终端的状态，如接收用户图像正在被播放、某终端加入或退出会议、终端号码分配等。

2. H.245 主要控制过程

H.245 控制信令协议的主要功能是建立和维护 RTP 会话，在 RTP 会话参与者之间协商所支持的功能，这些功能包括多媒体信息流编码格式、所需带宽等。H.245 控制信令协议通过能力交换、主从确定、往返时延确定、逻辑信道打开和关闭、环路维护、呼叫释放等协议控制过程完成上述功能。

（1）能力交换过程

能力交换过程是 H.225.0 呼叫建立成功后首先要执行的过程，它使通信双方了解对方接收和发送信号的能力。在 H.245 事务中，如果两个 H.323 终端意欲进行媒体通信，首先必须彼此了解对方的接收能力，也就是知道对方能够对那些格式的媒体数据进行解码。只有知己知彼，方能快速建立一个双方都能够接受的媒体信道。当然如果能够知道对方的编码能力会更好，这样它就可以挑选、要求对方发送自己喜欢的码流格式。所以，在 H.245 事务中，在建立媒体信道之前，必须首先完成能力交换。能力交换的过程将保证传输的多媒体信号能被接收并能得到相应处理。

在实际实现中，两个终端通过互发 TerminalCapabilitySet 消息来告知对方自己的编解码能力。完成能力交换后，某一端就可以根据对方的能力，请求建立一个或者几个合适的逻辑信道。在网络、系统资源允许和能力集已经声明的条件下，通信的双方之间可以同时存在多条媒体信道，例如，一路视频、两路音频等。

在完成第一次能力交换后，任何时候都可以再进行能力交换。通常利用再次的能力交换来反映能力集的改变。虽然也可以发送没有任何改变的能力集，但是没有任何的实际意义，所以如果没有非常重要的原因，不应该再发送那些根本没有任何变化的能力集。

(2) 主从确定过程

主从确定是为了解决两个 H.323 端点同时打开双向信道时的冲突问题，也用于会议通信中解决两个均含 MC 功能的端点的 MC 冲突问题。

H.245 事务中可能有如下的事情发生：参与通话的两个终端同时发起同一事件，这时有可能会引起对资源的竞争冲突。例如，两个终端同时发起建立具有相同参数的双向逻辑信道请求，由于两个双向逻辑信道的参数完全相同，只需要有一个同样的逻辑信道建立即可，这时就有竞争存在，H.245 用主从确定事务来解决这种冲突。根据终端类型和一个随机的状态判断数，按照 H.245 中提供的算法，确定那个终端具有高的优先级，即那个终端是 Master，那个终端是 Slave。当发生冲突时，Master 胜出，赢得控制权，具有高的优先级。

一个完整的主从确定事务完成的标志是：参与通话的两端都已经确认对端知道了主从确定的结果，并且这个结果是一个确定的结果，也就是分出了 Master 和 Slave。

终端在确认对端知道主从确定结果后，并且自己也知道主从确定结果，才能启动那些依赖于主从确定结果的 H.245 事务。

主从确定的另一个作用是会议通信中的 MC (Multipoint Controller) 仲裁。由于一个会议呼叫只能有一个 MC，如果两个参会的 H.323 实体都含有 MC，则必须通过一定的规则确定其中一个是主 MC。

(3) 往返时延确定过程

某些应用需要知道发送终端和接收终端之间的往返时延，往返时延确定过程提供了测量该时延的一种机制。方法很简单，只包含两个无参数的消息：时延测量请求和响应，发起请求方本地测量发收这两个消息之间的时间间隔得出该时延值，该过程也可用于检测远端终端是否在工作（如是否关机）。

(4) 逻辑信道的打开和关闭

在 H.323 协议中，媒体数据是通过媒体信道来传递的，逻辑信道信令过程定义了如何打开和关闭音频图像和数据的逻辑信道。完成能力交换和主从确定后，H.323 终端就可以发起媒体信道建立请求，如果对端有能力对未来通过请求信道所发送的媒体类型解码，并且同意这个请求，它就回复 OpenLogicalChannelAck 给信道发起者，同意建立这个信道，并且已经做好准备接收从这个信道传来的媒体数据；这时信道发起者就建立这个信道，并且通过这个信道发送数据给对端。

一个信道只能传递一种类型的媒体数据，可以通过设置有关的参数 (Associated SessionID) 在两个媒体信道之间建立关联，典型的应用是音视频唇音同步。

逻辑信道由音视频数据发送端发起建立，与主、被叫无关；同时音视频信道

只能由音视频数据的发送方来执行实际的关闭操作，接收端只能请求发送端关闭媒体信道。

在等待媒体数据接收方回应 `CloseLogicalChannelAck` 时，存在一个超时机制，这里讨论其中的一种超时情况：媒体数据发送方发送的 `CloseLogicalChannel`，由于网络原因，可能根本就没有被媒体数据接收方收到，这时就会出现超时，在标准中提供的超时机制中，对此时媒体数据接收方的行为没有作规定；在实现中，对这种情况接收方也可以不作任何事情，这是因为可以留给用户来处理。由于发送方在发送 `CloseLogicalChannel` 时，已经停止发送媒体数据，接收端就会在一段时间内接收不到媒体数据，这时接收端的终端往往会选择停止或者重启服务。

(5) 环路维护过程

环路维护是一个常规的维护过程，经由专用消息通知对方配合进行环路测试，还包含一个环路测试，结束命令消息，此过程对于网关来说是必备功能。

(6) 呼叫释放过程（结束会话）

通信的任何一方都可以发起呼叫释放，首先端点 A 停止在逻辑信道上发送消息，关闭所有逻辑信道，然后在 H.245 控制信道上向端点 B 发送“结束会话”命令消息，其后停止发送 H.245 消息。端点 B 收到上述消息后，关闭所有逻辑信道，向端点 A 回送“结束会话”命令消息，至此 H.245 控制信道关闭。

3. H.245 消息流程

图 4.9 为两个端点之间进行 H.245 协商的一个消息流程示意图，包括 H.245 的终端能力交换、主从确认、逻辑信道打开和最终的逻辑信道关闭、呼叫释放（结束会话）等主要控制过程。

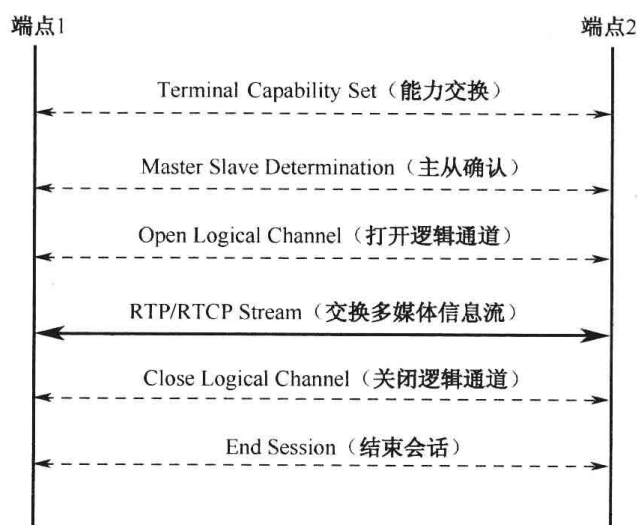


图 4.9 H.245 消息流程

4.1.5 H.323 呼叫流程实例

图 4.10 所示为在同一网守控制下的两个 H.323 端点用户之间的一个典型呼叫信令流程，包括 H.225.0 呼叫建立过程、H.245 控制信令的终端能力交换、主从确认、逻辑信道打开过程，RTP 媒体流传送以及 H.225.0 呼叫释放等过程。

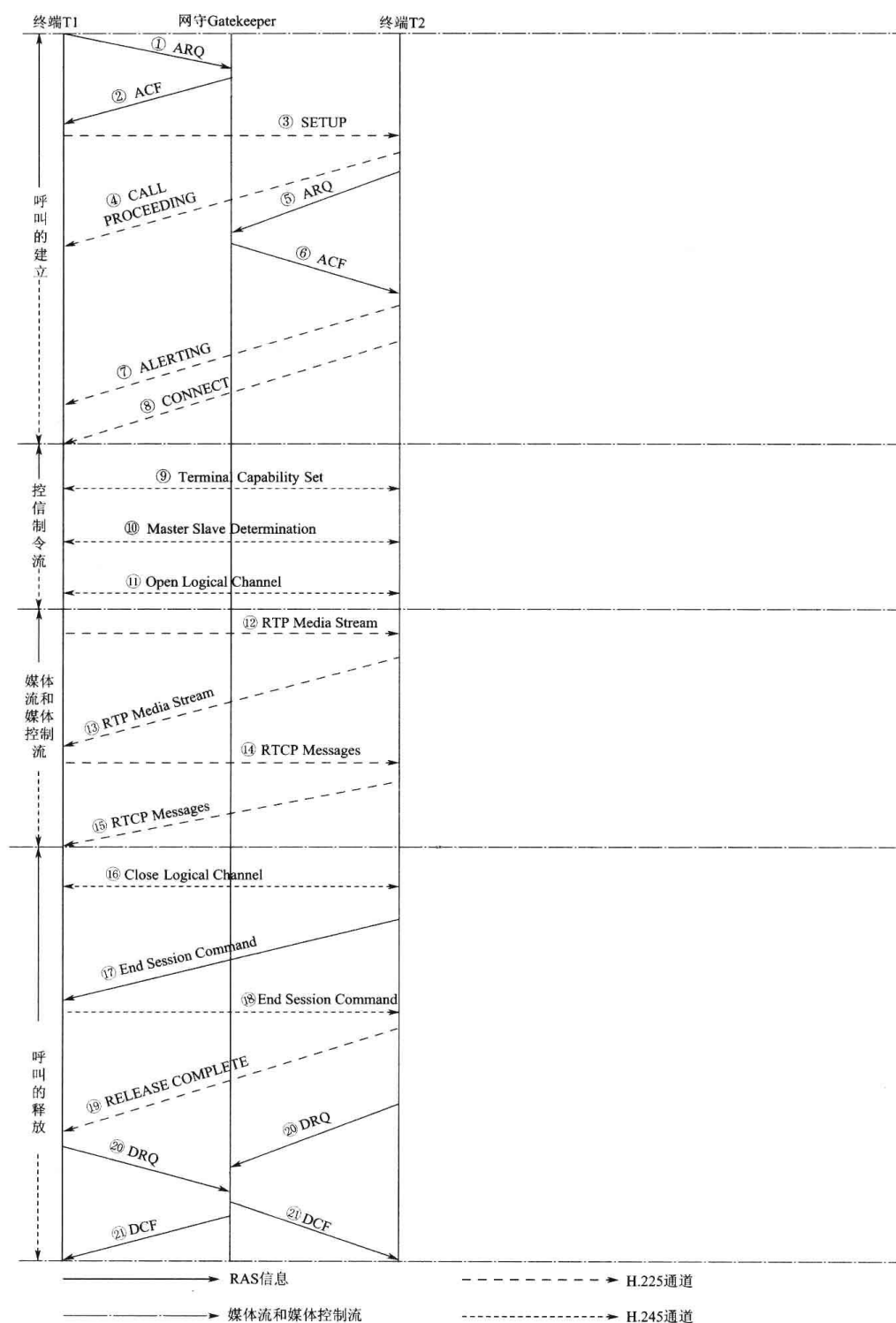


图 4.10 同一网守控制下的两个 H.323 端点用户之间的成功呼叫流程

具体流程如下所述。

① T1 在 RAS 通道上发送 RAS 请求信息到网守，请求使用 Direct Call Singaling 直接呼叫方式。

② 网守通过发送 ACF 信息到 T1 确认请求，允许 T1 使用 Direct Call Singaling 直接呼叫方式。

③ T1 发送一个 H.225 呼叫信令建立信息到 T2 请求连接。

④ T2 发送一个 H.225 呼叫处理信息回应 T1 的请求。

⑤ T2 在 RAS 通道上发送 RAS 请求信息到网守，要求注册本次呼叫。

⑥ 网守通过发送 ACF 信息到 T2 来确认注册。

⑦ T2 发送 H.225 警告信息给 T1，要求 T1 准备连接。

⑧ T2 发送 H.225 连接信息给 T1 确认建立连接，完成呼叫的建立。

⑨ T1 和 T2 之间已建立起 H.225 控制信道，开始能力交换过程，两个终端互发 TerminalCapabilitySet 消息来告知对方自己的能力。

⑩ T1 和 T2 之间进行主从确认。

⑪ T1 和 T2 之间的逻辑通道的打开过程通过一对 OpenLogicalChannel 消息和 OpenLogicalChannelAck 消息来完成。

⑫ T1 发送 RTP 封装的媒体流给 T2。

⑬ T2 发送 RTP 封装的媒体流给 T1。

⑭ T1 发送 RTCP 信息给 T2。

⑮ T2 发送 RTCP 信息给 T1。

⑯ 关闭逻辑通道过程。单向逻辑信道只能由发送者发起关闭逻辑通道过程，发送者发送一个 CloseLogicalChannel 消息给接收方，接收方返回一个 CloseLogicalChannelAck 来确认成功关闭了逻辑信道。

⑰ T2 开始呼叫的释放，它发送 H.245 终止会话命令给 T1。

⑱ T1 开始断开呼叫的终端，并发送 H.245 终止会话命令给 T2 来确认断开。

⑲ T2 发送一个 H.225 释放结束命令给 T1 来结束呼叫。

⑳ T1 和 T2 发送 RAS DRQ 信息给网守，使自己脱离网守。

㉑ 网守发送 DCF 信息给 T1 和 T2 来确认脱离。

4.2 基于 SIP 协议的视频会议网络技术

SIP (Session Initiation Protocol) 是一个应用层的信令控制协议。用于创建、修改和释放一个或多个参与者的会话。这些会话可以是 Internet 多媒体会议、IP 电话或多媒体分发。会话的参与者可以通过组播 (Multicast)、网状单播 (Unicast)

或两者的混合体进行通信。

使用 SIP，服务提供商可以随意选择标准组件，不论媒体内容和参与方数量，用户都可以查找和联系对方。SIP 对会话进行协商，以便所有参与方都能够就会话功能达成一致并进行修改，它甚至可以添加、删除或转移用户。

SIP 既不是会话描述协议，也不提供会议控制功能。为了描述消息内容的负载情况和特点，SIP 使用 Internet 会话描述协议（SDP）来描述终端设备的特点。SIP 自身也不提供服务质量（QoS），它与负责语音质量的资源预留协议（RSVP）互操作，还与若干个其他协议进行协作，包括负责定位的轻型目录访问协议（LDAP）、负责身份验证的远程身份验证拨入用户服务（RADIUS），以及负责实时传输的 RTP、RTCP 等多个协议。SIP 如同电子邮件客户机一样重用 MIME 类型描述，因此与会话相关的应用程序可以自动启动。SIP 重用几个现有的比较成熟的 Internet 服务和协议，如 DNS、RTP、RSVP 等。不必再引入新服务对 SIP 基础设施提供支持，因为该基础设施很多部分已经到位或现成可用。

SIP 独立于传输层。因此底层传输可以是采用 ATM 的 IP，SIP 使用用户数据报协议（UDP）和传输控制协议（TCP），将独立于底层基础设施的用户灵活地连接起来。

SIP 较为灵活，可扩展，而且是开放的，它激发了 Internet 以及固定和移动 IP 网络推出新一代服务的威力。SIP 能够在多台 PC 和电话上完成网络消息，模拟 Internet 建立会话。

这种可扩展性已经在越来越多基于 SIP 的服务中取得重大成功，本章主要从以下几个方面讲述基于 SIP 的视频会议技术。

- ① SIP 的逻辑实体；
- ② SIP 的基本消息；
- ③ SDP 协议概述；
- ④ RTP 协议概述；
- ⑤ RTCP 协议概述；
- ⑥ 基于 SIP 的视频会议系统。

4.2.1 SIP 的逻辑实体

SIP 被称为会话发起协议，是一个应用层的信令控制协议。会话的参与者可以通过组播（Multicast）、网状单播（Unicast）或两者的混合体进行通信。SIP 在建立和维持终止多媒体会话协议上，支持以下 4 种功能。

- 用户定位：检查终端用户的位置，用于通信；
- 用户有效性：检查媒体和媒体的参数；

- 建立会话：建立会话参数在呼叫方和被叫方；
- 会话管理：包括发送和终止会话，修改会话参数，激活服务等。

SIP 本身不提供服务，但是，SIP 提供了一些基本功能，可以用来实现不同的服务。比如，SIP 可以定位用户和传输一个封装好的对象到对方的当前位置。如果我们利用这功能，通过 SDP 传输会话的描述，对方的用户代理立刻就可以得到这个会话的参数。这个简单的例子说明，SIP 作为一个基础协议，可以在其上提供很多不同的服务。

SIP 协议定义了以下 4 种逻辑实体。

(1) 用户代理 (User Agent)

用户代理分为客户端 (User Agent Client) 和用户代理服务器 (User Agent Server) 两部分。在客户端，客户端应用程序初始 SIP 请求并负责发起呼叫；在用户代理服务器端，服务器应用程序负责接收呼叫并做出响应。

(2) 用户代理服务器 (Proxy Server)

用户代理服务器负责接收用户代理发来的请求，根据网络策略将请求发给相应的服务器，并根据收到的应答信息对用户做出响应。它可以转发客户端的请求，在转发之前可能修改原请求消息的内容。

(3) 重定向服务器 (Redirect Server)

重定向服务器接收 SIP 请求消息，并将被呼叫方的 SIP 地址映射成 0 个（如果没有可知地址）或更多的新地址，并且将它们返回客户端。

(4) 注册服务器 (Register Server)

注册服务器接收用户端的注册请求，完成用户地址的注册。

以上几种服务器可以共存于一个设备，也可以分布在不同的物理实体中。SIP 服务器完全由纯软件实现，可以根据需要运行于各种工作站或专用设备中。当各种逻辑实体分布在网络中时，用户之间通过这些实体的信令交互完成通信，具体流程如图 4.11 所示。首先用户代理向代理服务器发出 INVITE 请求，各个代理服务器之间进行消息转发，最终到达目的用户代理，如果这个用户代理同意接收消息，返回 200 OK，发出请求消息的用户代理在收到 200 OK 后，回送 ACK 确认，然后就可以完成媒体通信了。当通话结束时，任意一方发出 BYE 消息，接收方直接回送 200 OK。

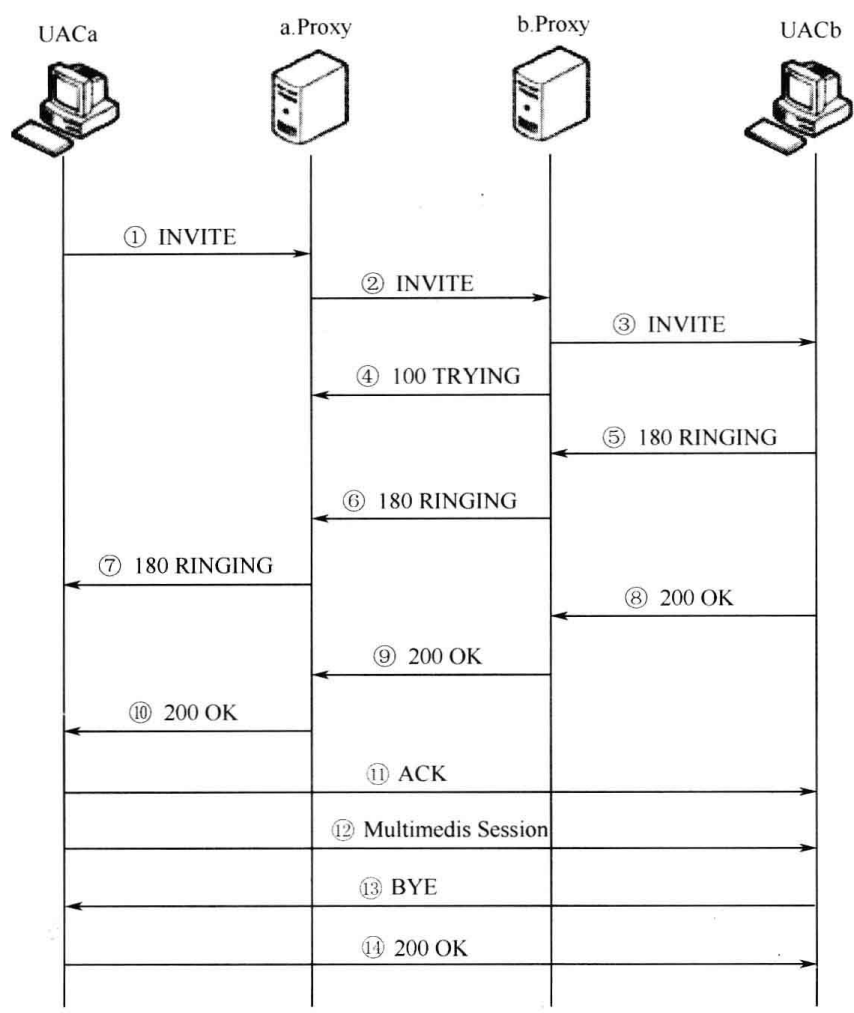


图 4.11 SIP 流程图

4.2.2 SIP 基本消息

SIP 包含下列基本消息，应符合 RFC3261 及相关扩展协议的规定，如表 4.5 所示。

表 4.5 SIP 消息描述

请求消息	
INVITE	业务请求消息，同时描述发送和希望接收的媒体类型
ACK	在收到对方对 INVITE 的最终响应消息后，向对方发送 ACK，表明接受对方对 INVITE 的响应
OPTION	用于询问对方能力集
BYE	表明释放呼叫，收到 BYE 的一方必须停止该呼叫中媒体流的收发
CANCEL	用于消除一个未完成的请求，常用于取消 INVITE 请求，一个典型的例子是主叫在拨打被叫号码后，对方尚未摘机，尚未对 INVITE 请求进行回应，主叫主动终止呼叫请求
REGISTER	向 SIP 服务器登记
INFO	用于在会话中携带会话控制信息

续表

响应消息	
1XX	1XX 信息消息，表明正在处理之中，还未能给出一个确定的答案，可以不发送或发送 1 个或多个 1XX 响应，1XX 消息没有重传机制，常用的消息有： ● 100 Trying——正在处理请求； ● 180 Ringing——正在振铃； ● 181 Call is being forwarded——Proxy 使用，表示呼叫转移到其他目的地； ● 182 Queued——被叫暂时不可达，正在排队等候处理。如果被叫变为有效，发送最终的响应
2XX	2XX 成功消息，表示请求成功，例如，200 OK 消息
3XX	3XX 重定向消息，用于告诉 UAC，将请求发送到一个新的地址
4XX	4XX 请求失败消息，常见的消息有： ● 400 Bad Request——不理解请求的内容； ● 401 Unauthorized——要求用户认证； ● 404 Not Found——在管理域内，请求的被叫用户不存在； ● 407 Proxy Authentication Required——与 401 相似，要求认证，不同之处是，401 用于注册认证，407 用于呼叫认证。例如，发送 INVITE 消息，UAS 返回 407，表示要求消息中带有用户身份认证的信息； ● 485 Ambiguous——请求中被叫的地址含糊或有歧义； ● 486 Busy here——成功到达用户的终端系统，但是用户暂时无效或不能接受呼叫
5XX	5XX 服务器错误消息
6XX	6XX 全局错误消息。例如，600 表示全忙

SIP 协议是一个基于文本的协议，使用 UTF-8 字符集。一个 SIP 消息既可以是一个从客户端到服务器端的请求，也可以是一个从服务器端到客户端的应答。即使在字符集上和语法细节上有所不同，请求和应答消息都基于 RFC2822 格式（SIP 允许包头域不是标准的 RFC2822 包头域）。这两种消息类型都由一个起始行、1 个或者多个包头域、一个可选的消息正文组成。

- 一般消息= 起始行；
- *消息包头；
- CRLF；
- [消息正文]；
- 起始行= 请求行 / 状态行；
- 起始行、包头行及空行都必须由回车换行组成，也必须有一个空行跟随。

1. SIP 请求消息

SIP 请求是根据起始行中的请求行来区分的。一个请求行包含方法名字、Request-URI，用单个空格（SP）间隔开。请求行由回车换行结束。除了用作行结束标志以外，不允许 CRLF 出现在其他地方。该规范规定了 6 种方法：REGISTER

用于登记联系信息；INVITE、ACK、CANCEL 用于建立会话；BYE 用于结束会话；OPTIONS 用于查询服务器负载，SIP 扩展、标准 RFC 追加可能包含扩展的方法。

- Request-URI: Request-URI 是一个 SIP 或者 SIPS-URI，也可以是一个通用的 URI (RFC2396)，它标志了这个请求所用到的用户或者服务的地址。Request-URI 禁止空白字符或者控制字符，并且禁止用“◇”括上。SIP 元素可以支持除了 SIP 或者 SIPS 之外所规定的 Request-URIS。
- 协议版本: 请求和应答消息都包含当前使用的 SIP 版本，一个应用发出的 SIP 消息一定包含了 SIP-Version。

2. SIP 响应消息

SIP 应答和 SIP 请求的区别在于在起始行中是否包含一个状态行。一个状态行在由数字表达的状态码之前，是一个协议的版本串，每一个元素之间用一个空格分开。状态码的第一个数字表示应答的类型。接下来两个数字并不作分类使用。基于这个原因，任何状态码 100~199 可以统称为“1XX 应答”，类似地，200~299 可以统称为“2XX 应答”，以此类推。SIP / 2.0 允许以下 6 类应答。

- 1XX: 临时应答——请求已经接收，正在处理这个请求。
- 2XX: 成功处理——请求已经成功接收，并且正确处理了这个请求。
- 3XX: 重定向——还需要附加的操作才能完成这个请求，本请求转发到其他服务器上处理。
- 4XX: 客户端错误——请求包含错误的格式或者不能在这个服务器上完成。
- 5XX: 服务器错误——服务器不能正确地处理这个显然合法的请求。
- 6XX: 全局故障——表示任何服务器无法完成该请求。

4.2.3 SDP 协议概述

SIP 协议仅仅完成了初始化连接，而媒体通信需要的信息，比如通信对方的视 / 音频编码格式、通信端口等信息是需要有专门的协议去沟通完成的。SIP 信令一般将 SDP 协议构建的内容作为消息体，因此，研究 SDP 协议是有必要的。

1. SDP 协议简介

会话描述协议 SDP 提供了统一的描述格式，包括会话信息和媒体信息，以实现会话者之间的能力协商。SDP 也是一种文本协议，采用 UTF.8 编码的 ISO10646 字符集。

SDP 分为会话描述和媒体级描述, 其中会话信息包括: 会话的名称和目的; 会话存活时间; 为接收媒体所需的信息 (地址、端口、格式等); 会话使用的带宽信息。媒体信息包括: 媒体类型 (音频、视频等)、传输协议 (RTP/UDP/IP、H.320 等)、媒体格式 (H.263 Video、MPEG Video 等); 多播或远端 (单播) 地址和端口。

下面是上述两种描述的基本含义。

1. 会话描述中几个重要文本行的格式和意义

① 协议版本: 格式如下。

v=0

② 会话源, 格式如下。

v=0 (用户名) <会话标识> (版本) (网络类型) <地址类型> (地址)

用户名是用户在发起主机上的登录名。会话标识为一数字串, 使得多元组 (用户名、会话标识、网络类型、地址类型、地址) 构成会话的全球唯一的标识符。会话标识的分配由创建工具决定, 一般采用网络时间协议 (NTP) 时戳, 以确保唯一性。网络类型为文本串, 目前规定 “IN” 表示 Internet。地址类型也是文本串。

③ 连接数据: 格式如下。

c^ (网络类型) (地址类型) (连接地址)

给出连接数据, 每个媒体描述必须包含一个 “c=” 字段, 或者在会话描述中包含一个公共的 “c-” 字段。

④ 属性: 格式如下。

a-- (属性) 或 a-- (属性): (值) 令 a= (flag)

表示特性属性, 它们为二进制属性, 标识会话具有该性质。

2. 媒体级描述的格式和意义

(1) m= (媒体类型) (端口) <传送层><格式列表>

① 媒体类型: 定义了 5 种类型, 包括音频、视频、应用、数据和控制。

② 端口: 指明媒体流发往的传送层端口, 其含义取决于 “c=” 行规定的网络类型和 “m=” 行定义的传送层协议。

③ 传送层协议: 其值和 “c-” 行中的地址类型有关。对于 IV 来说, 大多数媒体业务流都在 RTP/UDP 上传送。

④ 媒体格式: 对于音频和视频来说, 指 RTP 音频 / 视频应用文档中定义的媒

体净荷类型。

(2) SDP 格式

```

Session description
v=(protocol version)
o=(owner/creator and session identifier)
s=(session name)
i=(session information)
u=(URI of description)
e=(email address)
p=(phone number)
c=(connection information - not required if included in all media)
b=(zero or more bandwidth information lines)
One or more time descriptions("t=" and "r=" lines, see below)
z=(time zone adjustments)
k=(encryption key)
a=(zero or more session attribute lines)
Zero or more media descriptions
Time description
t=(time the session is active)
r=(zero or more repeat times)
Media description, if present
m= (media name and transport address)
i=(media title)
c=(connection information - optional if included at
  session-level)
b=(zero or more bandwidth information lines)
k=(encryption key)
a=(zero or more media attribute lines)

```

以上带*号的是可选的，其余的是必需的，一般顺序也按照上面的顺序来排列。

- a=*是 SDP 协议扩展属性定义，除上述内容外，分解时其他信息都可以扔掉。
- a=charset 属性指定协议使用的字符集，一般是 ISO-10646。

示例：

```
v=<username> <sess-id> <sess-version> <nettype> <addrtype> <unicast-address>
```

其中，nettype 是 IN，代表 internet，addrtype 是 IP4 或 IP6；unicast-address 代表任务创建计算机的地址。

该属性唯一表示一个任务：

e=123@126.com 或 p=+1 616 555-6011

对于一个具体任务只能是两者之中的一个，表示会议控制者的联系方式。邮件地址可以用[email]j.doe@example.com[/email] (Jane Doe) 形式表示，括号里面的信息描述联系人的名称，或者表示为 Jane Doe <[email]j.doe@example.com[/email]>，前面的信息是联系人的名称。

c=<nettype> <addrttype> <connection-address>这个连接数据可以是传话级别的连接数据，或者是单独一个媒体数据的连接数据。多播时，connection-address 就该是一个多播组地址；单播时，connection-address 就该是一个单播地址。对于 addrttype，在 IPv4 的情况下，connection-address 不仅包含 IP 地址，而且还要包含 a time to live value (TTL 0~255)，如 c=IN IP4 224.2.36.42/128。IPv6 没有这个 TTL 值，还允许<base multicast address>[/<ttl>]/<number of addresses>格式的 connection-address，如 c=IN IP4 224.2.1.1/127/3 等同于包含 c=IN IP4 224.2.1.1/127、c=IN IP4 224.2.1.2/127、c=IN IP4 224.2.1.3/127 三行内容。

b=<bwtype>:<bandwidth> bwtype 可以是 CT 或 AS，CT 方式设置整个会议的带宽，AS 方式设置单个会话的带宽，默认带宽单位是 kbps。

t=<start-time> <stop-time>可以有行，指定多个不规则时间段，如果是规则的时间段，则 r=属性可以使用。start-time 和 stop-time 都遵从 NTP (Network Time Protocol)，以秒为单位，自从 1900 年以来的时间，要转换为 UNIX 时间，减去 2 208 988 800。如果 stop-time 设置为 0，则会话没有时间限制。如果 start-time 也设置为 0，则会话被认为是永久的。

r=<repeat-interval> <active duration> <offsets from start-time>重复次数在时间表示里面可以如下表示：

```
d - days (86400 seconds)
h - hours (3600 seconds)
m - minutes (60 seconds)
s - seconds (allowed for completeness)
z=<adjustment time> <offset> <adjustment time> <offset> ....
k=<method>
k=<method>:<encryption key>
a=<attribute>
a=<attribute>:<value>
m=<media> <port> <proto> <fmt> ...
m=<media> <port>/<number of ports> <proto> <fmt> ...
```

其中，<media>可以是"audio"、"video"、"text"、"application"和"message"；<port>是媒体传送的端口号，它依赖于 c=和<proto>；<proto>可以是 udp、RTP/AVP 和 RTP/SAVP。

```

a=cat:<category> //分类, 根据分类接收者隔离相应的会话
a=keywds:<keywords> //关键字, 根据关键字隔离相应的会话
a=tool:<name and version of tool> //创建任务描述的工具的名称及版本号
a=ptime:<packet time> //在一个包里面的以毫秒为单位的媒体长度
a=maxptime:<maximum packet time> //以毫秒为单位, 能够压缩进一个包的媒体量
a=rtpmap:<payload type> <encoding name>/<clock rate> [<encoding parameters>]
a=recvonly
a=sendrecv
a=sendonly
a=inactive,
a=orient:<orientation> //其可能的值, "portrait", "landscape" and "seascape"
a=type:<conference type> //建议值是"broadcast", "meeting", "moderated", "test" and "H332"
a=charset:<character set>
a=sdplang:<language tag> //指定会话或者是媒体级别使用的语言
a=framerate:<frame rate> //设置最大视频帧速率
a=quality:<quality> //值为 0~10
a=fmtp:<format> <format specific parameters>

```

(3) 会话描述协议 (Session Description Protocol, SDP)

SDP 其实就是一个描述会话的格式标准, 又称协议, 大家都采用这种标准。就是把信息用 SDP 这种方式来格式化, 达到统一。下面描述了它的语法, 如何解析和填充相应字段, 必需有还是可选的, 在实际工程中需要细心选择。

具体格式如下:

```

Session description
v=(protocol version)
o=(owner/creator and session identifier)
s=(session name)
i=(session information)
u=(URI of description)
e=(email address)
p=(phone number)
c=(connection information - not required if included in all media)
b=(zero or more bandwidth information lines)
One or more time descriptions("t=" and "r=" lines, see below)
z=(time zone adjustments)
k=(encryption key)
a=(zero or more session attribute lines)
Zero or more media descriptions
Time description

```



```

t= (time the session is active)
r=*(zero or more repeat times)
Media description, if present
m= (media name and transport address)
i=*(media title)
c=*(connection information - optional if included at
    session-level)
b=*(zero or more bandwidth information lines)
k=*(encryption key)
a=*(zero or more media attribute lines)

```

以上带*号的是可选的，其余的是必需的，一般顺序也按照上面的顺序来排列。

- a=*是 SDP 协议扩展属性定义，除上述内容外，分解时其他信息都可以扔掉。
- a=charset 属性指定协议使用的字符集，一般是 ISO-10646。

示例：

```
v=<username> <sess-id> <sess-version> <nettype> <addrtype> <unicast-address>
```

其中，nettype 是 IN，代表 internet，addrtype 是 IP4 或 IP6；unicast-address 代表任务创建计算机的地址。

该属性唯一表示一个任务：

```
e=123@126.com 或 p=+1 616 555-6011
```

对于一个具体任务只能是两者之中的一个，表示会议控制者的联系方式。邮件地址可以用[*email*]j.doe@example.com[/*email*]（Jane Doe）形式表示，括号里面的信息描述联系人的名称，或者用 Jane Doe <[*email*]j.doe@example.com[/*email*]>表示，前面的信息是联系人的名称。

c=<nettype> <addrtype> <connection-address>这个连接数据，可以是传话级别的连接数据，或者是单独一个媒体数据的连接数据。多播时，connection-address 就该是一个多播组地址；单播时，connection-address 就该是一个单播地址。对于 addrtype，在 IPv4 的情况下，connection-address 不仅包含 IP 地址，而且还要包含 a time to live value（TTL 0-255），如 c=IN IP4 224.2.36.42/128。IPv6 没有这个 TTL 值，还允许<base multicast address>[/<ttl>]/<number of addresses>格式的 connection-address，如 c=IN IP4 224.2.1.1/127/3 等同于包含 c=IN IP4 224.2.1.1/127、c=IN IP4 224.2.1.2/127、c=IN IP4 224.2.1.3/127 三行内容。

b=<bwtype>:<bandwidth> bwtype 可以是 CT 或 AS，CT 方式设置整个会议的带宽，AS 方式设置单个会话的带宽，默认带宽单位是 kbps。

t=<start-time> <stop-time>可以有行，指定多个不规则时间段，如果是规则的

时间段, 则 `r=` 属性可以使用。start-time 和 stop-time 都遵从 NTP (Network Time Protocol), 以秒为单位, 自从 1900 年以来的时间。要转换为 UNIX 时间, 减去 2 208 988 800。如果 stop-time 设置为 0, 则会话没有时间限制; 如果 start-time 也设置为 0, 则会话被认为是永久的。

`r=<repeat-interval> <active duration> <offsets from start-time>` 重复次数在时间表示里面可以如下表示:

```
d - days(86400 seconds)
h - hours(3600 seconds)
m - minutes(60 seconds)
s - seconds(allowed for completeness)
z=<adjustment time> <offset> <adjustment time> <offset> ....
k=<method>
k=<method>:<encryption key>
a=<attribute>
a=<attribute>:<value>
m=<media> <port> <proto> <fmt> ...
m=<media> <port>/<number of ports> <proto> <fmt> ...
```

其中, `<media>` 可以是 "audio"、"video"、"text"、"application" 和 "message"; `<port>` 是媒体传送的端口号, 它依赖于 `c=` 和 `<proto>`; `<proto>` 可以是 udp, RTP/AVP 和 RTP/SAVP。

```
a=cat:<category> //分类, 根据分类接收者隔离相应的会话
a=keywds:<keywords> //关键字, 根据关键字隔离相应的会话
a=tool:<name and version of tool> //创建任务描述的工具的名称及版本号
a=ptime:<packet time> //在一个包里面的以毫秒为单位的媒体长度
a=maxptime:<maximum packet time> //以毫秒为单位, 能够压缩进一个包的媒体量。
a=rtpmap:<payload type> <encoding name>/<clock rate> [/<encoding parameters>]
a=recvonly
a=sendrecv
a=sendonly
a=inactive,
a=orient:<orientation> //其可能的值, "portrait", "landscape" and "seascape"
a=type:<conference type> //建议值是, "broadcast", "meeting", "moderated", "test" and "H332"
a=charset:<character set>
a=sdplang:<language tag> //指定会话或者是媒体级别使用的语言
a=framerate:<frame rate> //设置最大视频帧速率
a=quality:<quality> 值是 0~10
a=fmtp:<format> <format specific parameters>
```

当 SIP 协议包含的内容是 SDP 时, 应该把 Content-Type 设置成 application/sdp。

下面是举例说明。

```
v=0
o=- 1175827911 1175827911 IN IP4 127.0.0.1
s=<No title>
i=<No author>(c)2002
c=IN IP4 0.0.0.0
t=0 0
a=SdpplinVersion:1610641560
a=StreamCount:integer;2
a=control:*
a=Flags:integer;9
a=IsRealDataType:integer;1
a=Copyright:buffer;"KGMpIDIwMDIA"
a=ASMRuleBook:string;"#($Bandwidth >= 0),Stream0Bandwidth = 44100, Stream1Bandwidth =
180900;"
a=range:npt=0-7.779000
m=audio 0 RTP/AVP 101
b=AS:44
b=RR:1653
b=RS:551
a=control:streamid=0
a=range:npt=0-7.779000
a=length:npt=7.779000
a=rtpmap:101 x-pn-realaudio/1000
a=fmtp:101
a=mimetype:string;"audio/x-pn-realaudio"
a=Helix-Adaptation-Support:1
a=ActualPreroll:integer;1857
a=AvgBitRate:integer;44100
a=AvgPacketSize:integer;640
a=EndOneRuleEndAll:integer;1
a=EndTime:integer;7741
a=MaxBitRate:integer;44100
a=MaxPacketSize:integer;640
a=Preroll:integer;3714
a=OpaqueData:buffer;"LnJh/QAFAAAucmE1AAAAEAAFAAAATgAXAAACgAAAAAABQ
v+AAAAAAQAoAAgAAAAACsRAAArEQAAAAQAAJnZW5yY29vawEHAAAAAAQAQAAA
wgAACUArespond for options:
```


4.2.4 实时传输协议（RTP）

在视频会议中，用户之间通信需要相互传送视频数据包和音频数据包，在本文的设计中，传输层使用的协议是 RTP（Real Time Transport Protocol）。下面将简要地对相关概念进行介绍。

RTP 是针对 Internet 上多媒体数据流的一个传输协议，由 IETF（Internet 工程任务组）作为 RFC1889 发布。RTP 被定义为在一对一或一对多的传输情况下工作，其目的是提供时间信息和实现流同步。RTP 的典型应用建立在 UDP 上，但也可以在 TCP 或 ATM 等其他协议之上工作。RTP 本身只保证实时数据的传输，并不能为按顺序传送数据包提供可靠的传送机制，也不提供流量控制或拥塞控制，它依靠 RTCP 提供这些服务。

1. RTP 工作机制

威胁多媒体数据传输的一个尖锐的问题就是不可预料数据的到达时间。但是流媒体的传输需要数据适时地到达用以播放和回放。RTP 协议通过使用时间标签、序列号以及其他的结构来控制数据流实时到达。在流的概念中“时间标签”是最重要的信息，发送端依照即时采样在数据包里隐蔽地设置了时间标签。在接收端收到数据包后，就依照时间标签按照正确的速率恢复成原始的实时数据。不同的媒体格式的调时属性是不一样的。但是 RTP 本身并不负责同步，RTP 只是传输层协议，为了简化运输层处理，提高该层的效率，将部分运输层协议功能（如流量控制）上移到应用层完成。同步是应用层协议完成的任务。RTP 没有运输层协议的完整功能，不提供任何机制来保证实时地传输数据，不支持资源预留，也不保证服务质量。RTP 报文甚至不包括长度和报文边界的描述。同时 RTP 协议的数据报文和控制报文使用相邻的不同端口，这样大大提高了协议的灵活性和处理的简单性。

RTP 协议和 UDP 二者共同完成运输层协议功能。UDP 协议只传输数据包，不管数据包传输的时间顺序。RTP 协议的数据单元是用 UDP 分组来承载的。

在承载 RTP 数据包时，有时一帧数据被分割成几个具有相同的时间标签的包，由此可知，时间标签并不是必需的，而 UDP 的多路复用功能使得 RTP 协议利用支持显式的多点投递来满足多媒体会话的需求。

RTP 协议虽然是传输层协议，但是它没有作为 OSI 体系结构中单独的一层来实现。RTP 协议通常根据一个具体的应用来提供服务，RTP 只提供协议框架，开发者可以根据应用的具体要求对协议进行充分扩展。

2. 封包结构

RTP 包结构如图 4.12 所示。

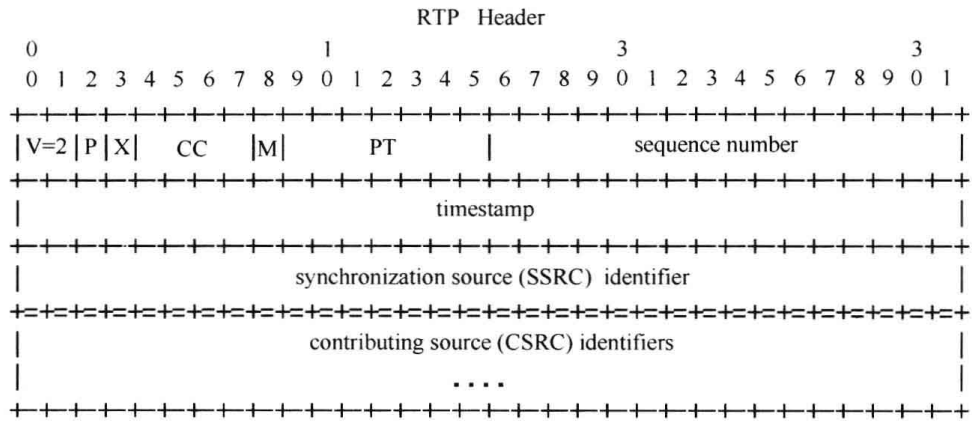


图 4.12 RTP 包结构

- V: 2 bit, 协定的版本号码, 目前版号是 2。
- P: 1 bit, RTP 封包 (Packet) 结束点的预留空间, 由封包是否需要多余的填充空间决定。
- X: 1 bit, 表明在封包中是否使用延伸空间。
- CC: 4 bit, 包含 CSRC 数目, 用于修正标头 (Fixed Header)。
- M: 1 bit, 用于定义应用等级及其原型 (Profile)。如果不为零, 则表示目前的资料有特别的程式解译。
- PT: 7 bit, 指 Payload 的格式并决定将如何由应用程式加以解译。
- SSRC: 同步化来源。

4.2.5 实时传输控制协议 (RTCP)

RTCP: 就是 RTP 控制协议 (RTP Control Protocol, RTCP)。

RTP 控制协议 (RTCP) 采用与数据包相同的分发机制, 将控制包周期性地传输到所有会话参与者中。底层协议必须提供数据和控制包的多路发送, 例如, 使用不同的 UDP 端口号。

RTCP 提供数据分发质量反馈信息, 这是 RTP 作为传输协议的部分功能, 涉及其他传输协议的流控制和拥塞控制。

RTCP 为 RTP 源携带一个持久性传输层标识符, 称为规范名或 CNAME。由于一旦发现冲突或程序重启时, SSRC 标识符会随之改变, 所以接收方需要用规范名来跟踪每一个参与者。同时, 接收方还要求规范名能够与一组相关 RTP 会话中来

自给定参与者的多重数据流相关联，例如，同步视频数据流和音频数据流。

上述前功能要求所有的参与者都要发送 RTCP 包，因此，必须控制速率以便 RTP 按比例增加大量的参与者。通过让每一个参与者发送各自的控制包给其他所有参与者，每一个参与者能够独立观察到参与者数量，该数量可用来计算控制包的发送速率。

OPTIONAL 功能用于传送最少会话控制信息，例如，在用户界面显示参与者标识。这对于“松散受控”会话（在没有成员控制或阐述协商的情况下，参与者可以加入或退出该会话）是非常有用的。

上述功能适用于所有环境，尤其是 IP组播环境。RTP 应用程序设计者应该避免设计只能工作于单播模式并且不能增加到大量数量的机制。在某些情况下，如在单向链接中，不可能有来自接收方的反馈，所以 RTCP 的传输就可能由发送方和接收方分别独立控制。

1. 协议结构

(1) 定义

单位：bit

Version	P	RC	Packet
2	3	8	16

(2) 数据包大小 (Length)

- Version——识别 RTP 版本。RTP数据包中的该值与 RTCP 数据包中的一样，当前规范定义的版本值为 2。
- P——间隙 (Padding)。在设置时，RTCP 数据包包含一些其他间隙八位位组，它们不属于控制信息；间隙的最后八位用于计算应该忽略多少个间隙八位位组，在一些加密算法中，当需要计算固定块大小时可能需要使用间隙字段。在一个复合 RTCP 数据包中，只有最后的个别数据包中才需要使用间隙，这是因为复合数据包是作为一个整体来加密的。
- RC——接收方报告计数，包含在该数据包中的接收方报告块的数量，有效值为 0。
- Packet Type——包括常量 200，识别这是一个 RTCP 源数据包。
- Length——RTCP数据包的大小（32 位字减去 1），包含头和任意间隙（偏移量的引入使得 0 成为有效值，并避免了扫描复合 RTCP 数据包过程中的无限循环现象，而采用 32 位字计数方法则避免了对 4 的倍数的有效性校验）。

2. RTCP 的分组类型

RTCP 5 种分组类型如表 4.6 所示。

表 4.6 RTCP 5 种分组类型

类型	缩写表示	意义
200	SR	发送端报告
201	RR	接收端报告
202	SDES	远点
203	BYE	结束
204	APP	特定应用

(1) 结束分组 (BYE)

结束分组 (BYE) 表示关闭一个数据流。

(2) 特定应用分组 (APP)

特定应用分组 APP 使应用程序定义新的分组类型。

(3) 接收端报告分组 (RR)

接收端报告分组 (RR) 用来使接收端周期性地向所有的点用多播方式进行报告。接收端每收到一个 RTP 流 (一次会话包含多个 RTP 流) 就产生一个接收端报告分组 RR。RR 分组的内容包括:

- 所收到的 RTP 流的同步源标识符;
- 该 RTP 流的分组丢失率 (若分组丢失率太高, 发送端就应该适当地降低发送分组的速率);
- 在该 RTP 流中的最后一个 RTP 分组的序号;
- 分组到达时间间隔的抖动等。

发送接收端报告分组有两个目的, 第一, 可以使所有的接收端和发送端了解当前网络的状态; 第二, 可以使所有发送 RTCP 分组的站点自适应地调整自己发送 RTCP 分组的速率, 使得起控制作用的 RTCP 分组不要过多地影响传送应用数据的 RTP 分组在网络中的传输。通常的做法是使 RTCP 分组的通信量不超过网络中数据分组数据量的 5%, 而接收端的通信量又应小于所有 RTCP 分组的通信量的 75%。

(4) 发送端报告分组 (SR)

发送端报告分组 (SR) 用来使发送端周期性地向所有接收端用多播方式进行

报告。发送端每发送一个 RTP 流就要发送一个发送端报告分组 SR。SR 分组的内容包括:

- 该 RTP 流的同步源标识符;
- 该 RTP 流中最新产生的 RTP 分组的时间戳和绝对时钟时间或用户时钟时间 (Wall Clock Time);
- 该 RTP 流包含的分组数;
- 该 RTP 流包含的字节数。

绝对时钟时间是必要的, 因为 RTP 要求每一种媒体使用一个流。例如, 要传送视频图像和相应的声音就需要传送两个流。有了绝对时钟时间就可以实现图像和声音同步。

(5) 源点描述分组 (SDES)

源点描述分组 (SDES) 给出对会话中参加者的描述, 它包含参加者的规范名 CNAME (Canonical NAME)。规范名是参加者的电子邮件地址的字符串。

3. RTCP 的实现

(1) 概述

一个 RTCP 流包含数据包格式、定时规则、参与者数据库三部分。

所有的 RTCP 复合包被周期性送出, 该周期被称为报告间隔。除了更新源描述和唇同步信息, 以及在该时间间隔内发生的接收质量统计操作以外, 所有的 RTCP 活动都在该间隔内发生。

参与者数据库是基于收到的 RTCP 包建立的, 其作用如下:

- ① 根据该数据库填充接收报告, 并发送给对方;
- ② 维护参与者信息;
- ③ 用于进行唇同步。

(2) RTCP 的传输

- ① 必须发送 RTCP 复合包;
- ② 目标端口是 RTP 端口号+1 (最近不要求必须是奇数, 也不要求必须大 1);
- ③ 所有的参与者应当送出复合包, 同时接收所有其他的参与者发送的复合包。

注意: 反馈给所有参与者的是一个多方会话, 可以单播到转发器, 然后重新分配数据; 或者直接通过多播完成。点对点的性质使每个参与者在进行流媒体会话时能获取到其他所有参与者 (如他们的存在) 的信息——接收质量和可选的个人详细信息, 如姓名、电子邮件地址、位置和电话号码等。

(3) RTCP 的包格式

RTCP 共有 5 种报文，包括 SR、RR、SDES、BYE、APP，其通用包格式如图 4.13 所示。

RTCP头部的格式				
V=2	P	RC	PT=SR=200	Length
SSRC of sender				
NTP timestamp, most significant word				
NTP timestamp, least significant word				
RTP timestamp				
Sender's packet count				
Sender's octet count				
SSRC_1(SSRC of first source)				
Fraction lost		Cumulative number of packets lost		
Extended highest sequence number received				
Interarrival jitter				
Last SR (LSR)				
Delay since last SR (DLSR)				
SSRC_2(SSRC of second source)				
...				
Profile-specific extensions				

图 4.13 SR、RR、SDES、BYE、APP 通用包格式

- V: 2 bit，协定的版本号码，目前版号是 2。
- P: 1 bit，RTP 封包（Packet）结束点的预留空间，由封包是否需要多余的填充空间决定。
- X: 1 bit，表明在封包中是否使用延伸空间。
- RC: 4 bit，包含 CSRC 数目，用于修正标头（Fixed Header）。
- PT: 7 bit，指 Payload 的格式并决定将如何由应用程式加以解译。
- Length: 16 Byte，报文长度。

1) RR(Receiver Report)

① RR 的作用：

接收质量报告提供所有发送 RTP 数据的发送方的信息，每个时隙包含一个 SSRC 的 RTP 接收质量报告。

```
PT = 201
Format:
Reporter SSRC
*{//一个 Reporter Block
固定头，24 octets 的内容，包括以下部分。
reportee SSRC:
cumulative number of packets lost //24 bit 的有符号数，从会话开始到现在期望收到——
实际收到（可为负）
extended highest sequence number:per session
loss fraction :per interval //取整（丢包 / 期望收到数目×256）（如果丢包为负
```


值, 则结果设为 0)

interarrival jitter:

last sender report timestamp(LSR) //从 reportee 端最后收到的 Sender Report 中 NTP timestamp 的中 32 bit (无则为 0)

delay since last sender report(DLSR) //最后收到 SR 和发送 RR 之间的间隔, 以 1/65536 为单位 (否则为 0)

}

② RR 的解释

RR 是对接收质量的反馈, 对发送方和其他的第三方如网络的监控等都有意义, 例如:

- 可以计算往返时间 (RR 收到时间-LSR-DLSR)。往返时间的估计是很重要的。
- 丢失: 短期内, 对媒体格式的选择有参考。
- 抖动: 突然增大的抖动通常意味着丢包的开始。

2) SR (Sender Report)

① SR 的作用

SR 提供有关发送方所发送的媒体的信息, 主要用于进行多个媒体流同步等。

PT = 200

Format:

固定头

Reporter SSRC

NTP Timestamp (2 octets): 发送 SR 的 NTP

RTP Timestamp: 与 previous field 同一时刻, 但是单位是 RTP Media clock

Sender's packet count: 自会话开始

Sender's octet count: 自会话开始

② SR 的解释

SR 是发送方发送的报告, 对于实现发送方和其他第三方的时钟同步有意义, 其作用如下:

- 计算速率等;
- 提供媒体时钟和 NTP 时钟。

3) SDES: 源描述 (Source Description)

用于提供参与者的详细信息, 通常是可读的, 如 E-mail、电话等, 依赖于应用。

PT = 202

Format:

固定头

*{

SSRC/CSRC

```
List of SDES items
```

```
}
```

其中每部分说明如下：

- Type Length content: 8 8。
- Type = 0 表示 Lists 结束。
- RFC 中规定了一些 Items，如 CNAME、NAME、EMAIL、PHONE、LOC、TOOL、NOTE、PRIV。

4) RTCP BYE: 成员控制

通常用于指示会话中的某个成员正在离开。但是 RTCP BYE 的含义在某种程度上依赖于应用，应用通常有其他的控制协议，如 SIP、H.323 等控制会话。

注意：BYE 不终止会话成员的关系，只是表示正在离开，并且不保证传输肯定到达，其长度和原因均可选。

```
PT=203
```

```
Format:
```

```
固定头
```

```
*{
```

```
SSRC
```

```
}
```

5) RTCP APP: Application-Defined RTCP Packets

Application-Defined RTCP Packets: 流媒体应用程序定义的数据包。

```
PT=204
```

6) Compound Packet

以上的 RTCP 包从不单独发送，它们被打包成复合包（Compound Packet）来发送，遵循以下几个规则：

- 对活动的发送者（Active Data Sender），Compound 必须以 SR 开始，否则以 RR 开始，即使没有数据接收和发送。后面跟着*个 RR。
- SDES、SDES 必须包含一个 CNAME，其他可选。
- 如果有 BYE 的话，放到最后，其他的包可以随意。

RTCP 数据包从来不单独传输，它们按照一定的规则总是组成复合包来传送，其规则是：最前面可以有个可选的 32 位值（在 RTCP 复合包加密时使用）。

（4）隐私与安全

SDES 的传输中会暴露隐私，这是个问题。但是规范名是必需的。

（5）包验证（Packet Validation）

- ① 所有的包必须是复合包。

- ② 版本必须是 2。
- ③ 复合包开始的 RTCP 包必须是 SR 和 RR。
- ④ 如果需要填充, 则只有最后一个数据包是可以填充的。
- ⑤ 所有的 RTCP 数据包的长度必须等于复合 RTCP 包的长度。

(6) 参与者数据库

该数据库用于存储参与者和会话的信息。

1) RTCP 的全局配置信息

- ① The RTP bandwidth: RTP 带宽。
- ② RTCP 所占总带宽的比例 (这意味必须知道 RTP 所占的总带宽): 默认时为 0.05。
- ③ 发送间隔: 默认时为 5 s (最小)。
- ④ 发送部分所占的比例: 默认时为 0.025。
- ⑤ 所有参与者收到/发送包的平均大小。
- ⑥ 参与者数量, 包括最后发送 RTCP 分组包的成员和会议过程中送 RTP 数据包包的成员。
- ⑦ 发送 RTCP 包时间与下一个预定发送的时间的间隔。
- ⑧ 发送的最后两个 RTCP 包——标记是否发出任何 RTP 数据包。
- ⑨ 标记是否发送所有的 RTCP 数据包。
- ⑩ 已经发送的 RTP 数据包的字节数。
- ⑪ 使用的最后一个序列号。
- ⑫ RTP 时钟与 NTP 时间戳格式之间的对应关系。

2) 会话中每个成员的信息

- ① 同步源识别。
- ② 源信息描述: CNAME 是必需的, 也可能包括其他信息 (注意: 其他信息不是以 NULL 结束的, 它们必须处理)。
- ③ 接收质量统计 (包括数据包的丢失和抖动), 允许生成 RTCP RR 包。
- ④ 从发送方收到的报告信息, 实现同步。
- ⑤ 不活跃的参与者接收最终消息时允许超时。
- ⑥ 指示该参与者已发送的数据在当前的 RTCP 报告间隔的标记。
- ⑦ 媒体播放缓冲区和任何所需的编码器状态 (媒体捕捉、播放时间)。
- ⑧ 任何信息的信道编码, 例如, 错误恢复所需的数据包的接收, 待修复后方可发送。

(7) 定时规则 (Time Rule)

总的目标是限制 RTCP 占 RTP 会话量的一小部分, 通常是 5%。在这个目标的

前提下,参与者送出 RTCP 数据包的速率是不固定的,而是变化的,如对有大量听众的网络广播,客户端可能几分钟才发送,而对双人对话的电话可能是几秒。

(8) 参考规则

有一些规则可以参考,这些规则可以确保实现的可升级性 (Scalability)。

1) 报告间隔

根据以下几个部分来计算:

- ① 流媒体平均带宽:通常是会话带宽的 5%。
- ② 发送和接收媒体数据包的平均大小:包括所有头 (UDP、IP)。
- ③ 所有参与者的数量和活跃者的数量。

计算过程如下所述,其中,

- SNo: 发送数
- ANo: 所有参与者数目
- RNo: 接收数
- Intl: 间隔
- ARS: RTCP 的平均大小
- RW: RTCP 流带宽

```
if (senders > 0 && senders <= (25% of total number of participants))
{
    if (we are sending)
    {
        Interval = average RTCP size * senders / (25% of RTCP bandwidth)
    }
    else
    {
        Interval = average RTCP size * receivers / (75% of RTCP bandwidth)
    }
}
else if ((senders = 0) or (senders > (25% of total number of participants)))
{
    Interval = average RTCP size * total number of members / RTCP bandwidth
}
//这样可以确保即使在发送者很少的情况下也可以确保 25%的带宽,能够送出足够的唇同步信息
If (Interval < minimum interval)
{
    Interval = minimum interval
}
```


当会话中的参与者的数目或者发送者所占比例发生变化时，报告间隔应当重新计算。

2) 基本的传输规则

基本的传输规则如下所述。

```
I = (Interval * random[0.5,1.5])
if (this is the first RTCP packet we are sending) {
  I *= 0.5
}
next_rtcp_send_time = current_time + I
```

对于有很多参与者并且数目还在变化的会话，每次发送当前的 RTCP 数据包后，都要根据得到的参与者数量来计算报告间隔。

3) 重发机制

当会话中同时有大量参与者时，会造成网络拥塞（Called as "step join"），为此需要使用重发机制。

方法：在每次发送前，根据当前的会话信息重新计算发送时间。

```
if (current_time >= next_rtcp_check_time) { //1.21828 是一个补偿因子
  new_rtcp_send_time = (rtcp_interval() / 1.21828) + last_rtcp_send_time
  if (current_time >= new_rtcp_send_time) {
    send RTCP packet
    next_rtcp_check_time = (rtcp_interval() / 1.21828) + current_time
  } else {
    next_rtcp_check_time = new_send_time
  }
}
```

4) 反向重发机制

当大量的人同时离开时，导致 RTCP 所占带宽过低（Called as "step leave"）。为此，当 BYE 消息到来时，需要立即重新计算下个包的发送时间。

5) BYE 重发机制

当存在大量的 BYE 消息时，可以延迟 BYE 的发送，或者干脆不发送。

6) 评论复议

各个实现应该考虑这个问题，并尽可能的使用各个复议。

7) 实现中常见的问题

- ① 没有正确的考虑参与者的数目，固定的发送间隔会导致流量呈线性增长。
- ② 发送间隔没有随机化。
- ③ 在带宽计算时未考虑报头（IP、UDP）等。
- ④ 填充使用不正确。

4.2.6 基于 SIP 协议的视频会议系统

业务组网结构如图 4.14 所示。

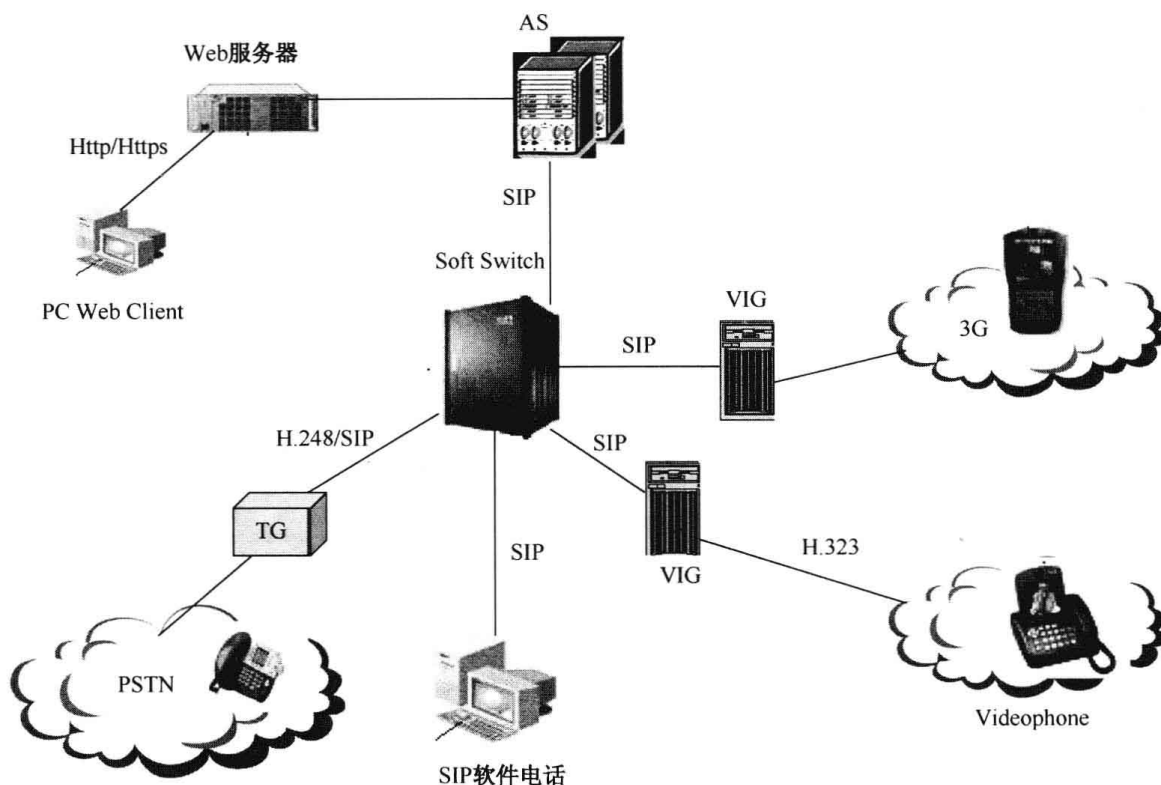


图 4.14 业务组网结构

多媒体视频会议系统主要由会议应用服务器、会议系统 Web 服务器组成，二者可以是同一个物理设备，也可以是独立的物理设备，MCU 也是会议系统的一部分，但是通常和会议应用服务器一起作为一个实体对待。会议服务器和软交换系统直接以标准 SIP 协议进行交互；PSTN 网络终端、H.323 网络终端和 3G 网络终端则通过 VIG 进行信令转化后以 SIP 协议接入软交换；如果软交换支持 H.323，则 H.323 网络可以直接接入软交换系统；SIP 软件电话则直接注册在软交换系统或者通过网络穿越设备注册在软交换系统中，进而使用会议应用服务器提供的业务。

用户通过 PC 中提供的 Web 浏览器访问 Web 服务器，进行会议管理，或者在会议中使用数据业务。

1. 系统功能描述

会议服务器借助宽带接入网络，不仅支持传统的语音会议，同时也支持视频会议和 Web 数据协作等多媒体业务。这样不但可以扩展传统的语音会议业务，同

时还可以推进宽带化、可视化的多媒体业务部署。

用户通过 Web 门户，建立音频、视频和数据会议。通过普通电话、IP 电话、视频终端、计算机和手机等各种终端参加会议。与传统的会议电视业务相比，会议服务器大大降低了系统的网络投资成本和终端成本。

2. 基本业务流程

基本业务流程如图 4.15 所示。

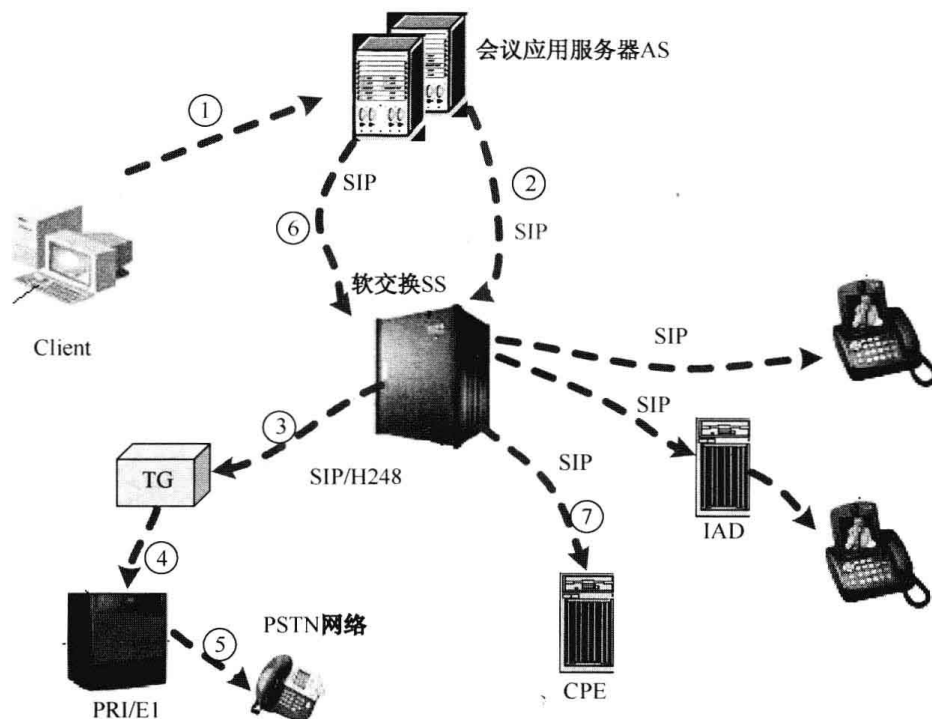


图 4.15 基本业务流程

下面以某客户通过多媒体会议系统实现多媒体会议为例，说明业务流程。

- ① 会议建立者通过 Web 在会议服务器上建立会议空间。
- ② 会议服务器通过 SIP 触发软交换呼叫会议参与者 A。
- ③ 软交换经过用户属性查找，此人为 PSTN 用户，将呼叫路由到 TG 上。
- ④ TG 通过 ISDN PRI 呼叫 PSTN 上的用户 A 的模拟话机。
- ⑤ 用户电话振铃，参与者 A 摘机，并进入会议系统。
- ⑥ 会议服务器继续呼叫第二个参与者，触发软交换呼叫参与者 B。
- ⑦ 软交换经过用户属性查找，此人为软交换下挂的直接用户；触发参与者 B 的 CPE 终端设备；参与者应答呼叫并进入会议系统。

3. 基本信令流程

(1) 创建会议（如图 4.16 所示）

- User A 通过 Web 在会议服务器上创建会议。
- 会议服务器返回创建成功或者失败的信息。

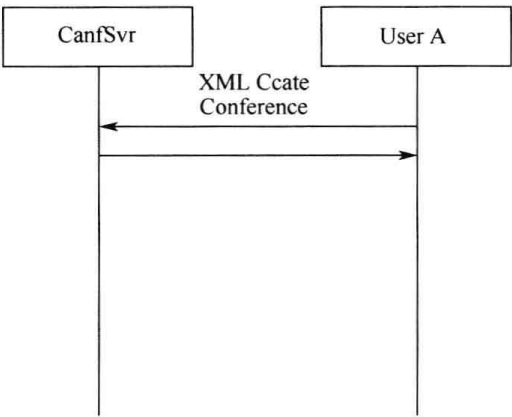


图 4.16 创建会议

(2) 会议通知

会议预定者可以向会议参与者发邮件或者短信通知会议信息，可以在会议服务器集成邮件系统来发送邮件，也可以使用短信模块或者短信网关来发送短信。

(3) 加入会议

1) 主动拨入会议（如图 4.17 所示）

USER A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。USER A 可以通过 Web 以邮件方式通知 User B 参加会议，并告知会议拨入号和参与者接入码。

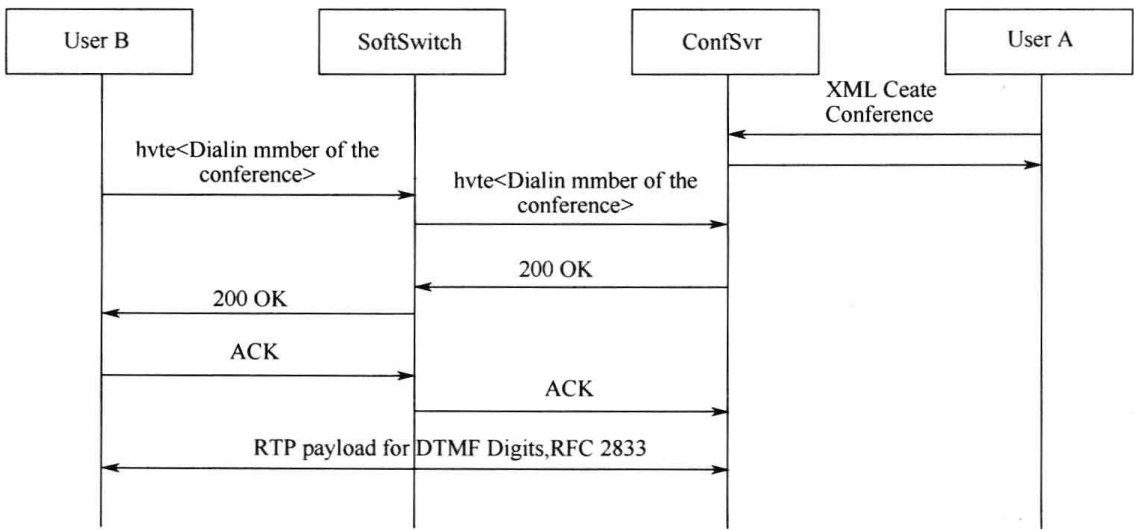


图 4.17 主动拨入会议

User B 通过会议终端接入会议接入号，通过软交换将接入会议请求传递到会议服务器，会议服务器将要求参与者输入接入码。如果会议终端是 SIP 电话，那么会议终端通过 SIP INVITE 请求拨入会议。

User B 通过会议终端接入会议接入号码，然后输入会议号或者会议密码即可参加会议。

① 音 / 视频分别邀请加入。

音 / 视频分别邀请加入如图 4.18 所示，具体流程如下所述。

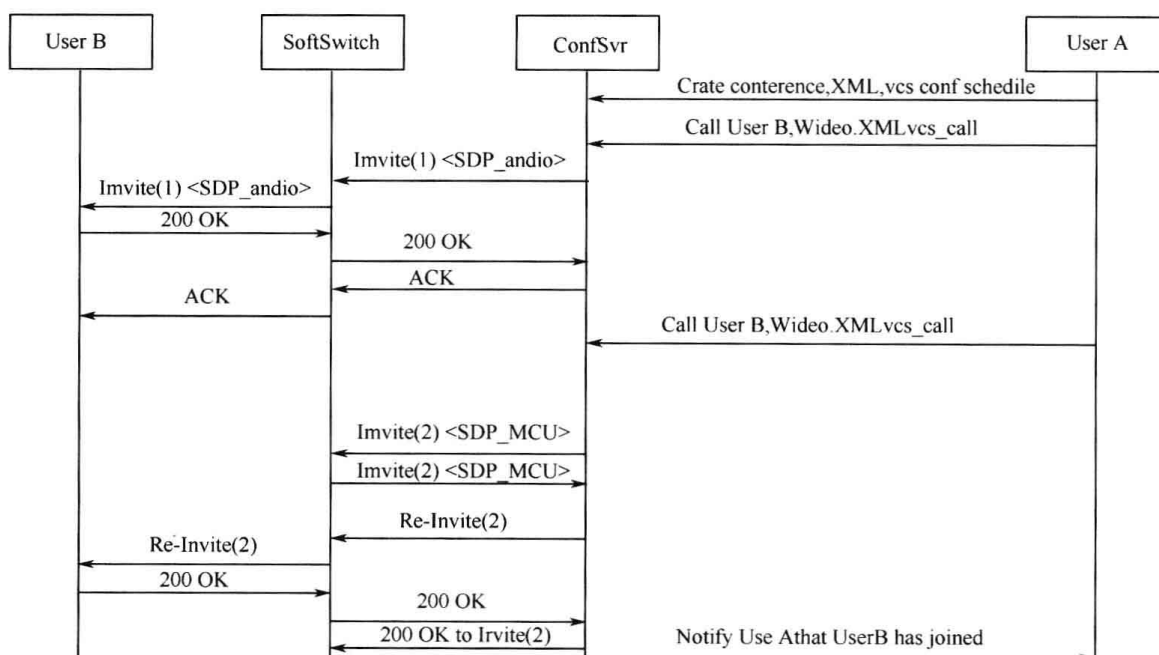


图 4.18 音 / 视频分别邀请加入流程

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。
- User A 可以通过 Web 先邀请 User B 参加音频会议。
- 会议服务器通过 SIP INVITE 请求触发软交换呼叫 User B 参加音频会议。
- User B 应答并加入会议。
- 如果 User B 的会议终端支持视频，那么 User A 可以通过 Web 再邀请 User B 加入视频会议。
- 会议服务器通过 SIP INVITE 请求触发软交换呼叫 User B 参加视频会议。
- 软交换通过 SIP re-INVITE 再次呼叫 User B 参加视频会议。
- User B 应答并加入会议。

② 音 / 视频一次邀请加入。

音 / 视频一次邀请加入如图 4.19 所示，具体流程如下所述。

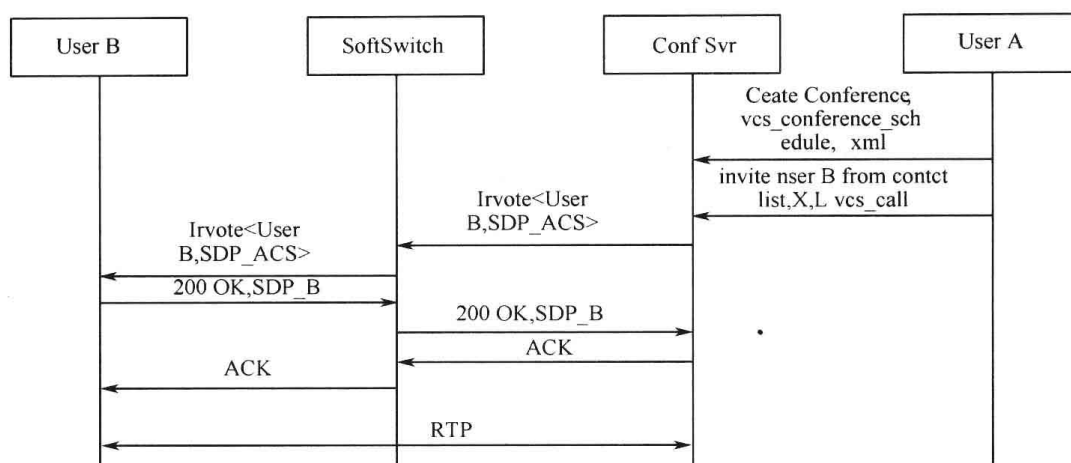


图 4.19 音 / 视频一次邀请加入流程

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。
- User A 可以通过 Web 邀请 User B 参加音 / 视频会议。
- 会议服务器通过 SIP INVITE 请求触发软交换呼叫 User B 参加音 / 视频视频会议。
- User B 应答加入会议。
- 如果 User B 的会议终端同时支持音视频，则音 / 视频都会接通，如果没有视频则只接通音频。

(4) 与会者退出会议

与会者退出会议如图 4.20 所示，具体流程如下所述。

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。
- User B 加入会议。
- User B 可以主动退出会议，通过挂机操作可以退出会议。如果 User B 使用 SIP 终端，那么 SIP BYE 请求消息将退出请求发送至软交换，软交换再将请求转发到会议服务器。

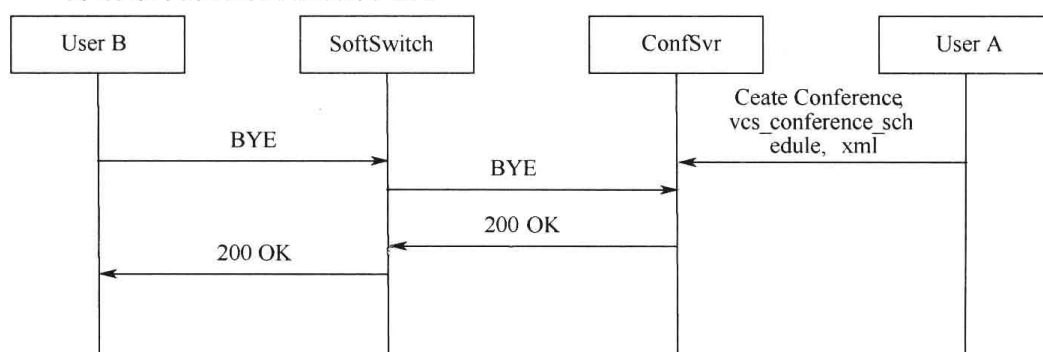


图 4.20 与会者退出会议

(5) 会议控制

1) 会议主持人增加会议成员

该流程和通过 Web 被动邀请加入会议的流程相同。

① 会议主持人通过 Web 将参与者退出会议。

会议主持人通过 Web 将参与者退出会议如图 4.21 所示，具体流程如下所述。

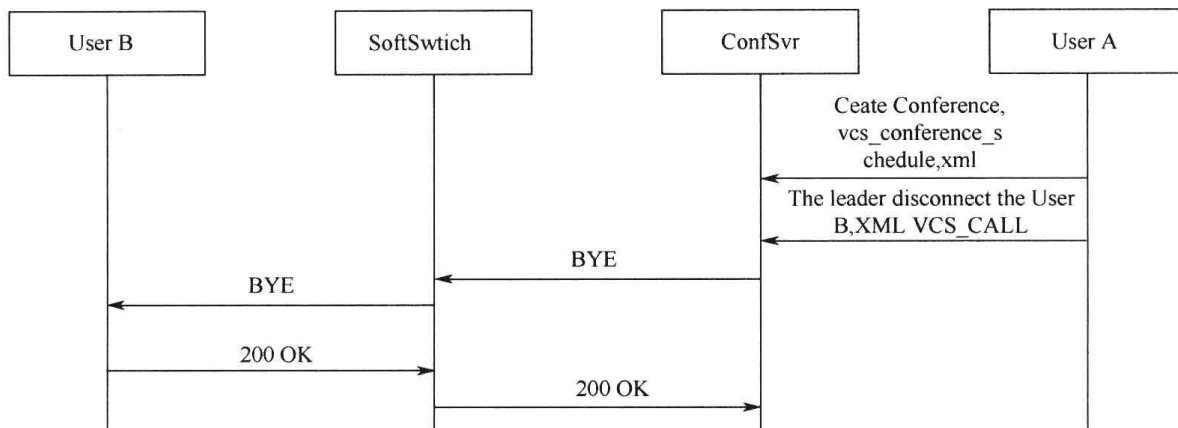


图 4.21 会议主持人通过 Web 将参与者退出会议

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。
- User B 加入会议。
- 会议主持人（如 User A）可以通过 Web 界面操作，将参与者 User B 退出会议。
- 会议服务器通过 SIP BYE 请求触发软交换断开 User B。
- 如果 User B 使用 SIP 终端，软交换转发 SIP BYE 请求断开 User B。

② 会议主持人锁定会议室。

会议主持人锁定会议室如图 4.22 所示，具体流程如下所述。

- 会议主持人如 User A 可以通过 Web 将会议锁定。
- 会议锁定后，User B 拨入会议系统，系统将会播放通告，告知会议已经锁定。
- 会议服务器断开与 User B 的通话。
- 会议主持人控制会议的发言权。
- 会议主持人 User A 可以通过 Web 将 User B 禁言。
- 会议服务器将不会在 RTP 中对 User B 的声音进行混音。
- 会议主持人 User A 可以通过 Web 将 User B 解禁言。
- 会议服务器将继续对 User B 的声音进行混音。

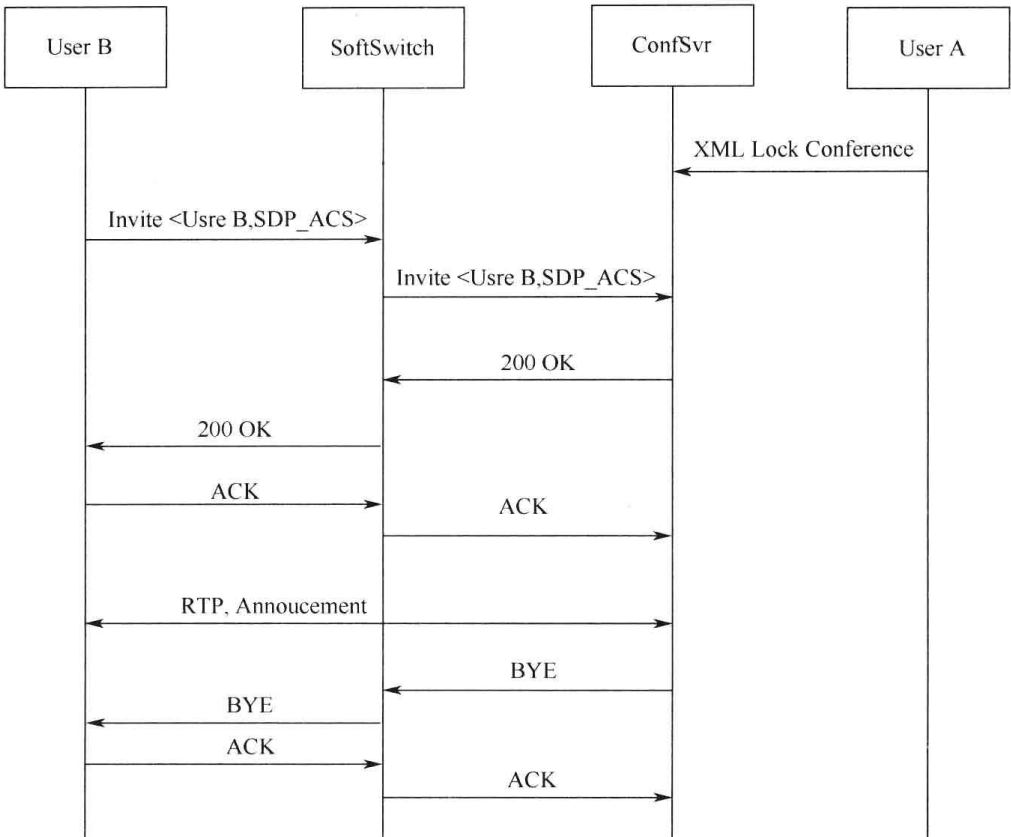


图 4.22 会议主持人锁定会议室

③ 会议主持人控制权转移。

- 会议主持人 User A 可以通过 Web 将会议管理权转给 User B 代理。

2) 会议参加者通告

会议参加者通告如图 4.23 所示，具体流程如下所述。

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。User A 可以指定会议参数，要求参与会议者参与会议前报出姓名通告其他与会者。User A 可以通过 Web 以邮件方式通知 User B 参加会议，并告知会议拨入号和参与者接入码。
- User B 通过会议终端拨叫会议拨入号，通过软交换将拨入会议请求传递到会议服务器，会议服务器将要求参与者拨入接入码。如果会议终端是 SIP 电话，那么会议终端通过 SIP INVITE 请求拨入会议。
- 会议服务器将要求参与者报出姓名以通告其他与会者。
- User B 报出姓名，其他与会者收听到通告。

3) 会议结束

会议结束如图 4.24 所示，具体流程如下所述。

- 在会议时间到时终止会议或者会议主持人主动终止会议。
- 会议服务器向会议参与人发送 BYE 消息，结束会话。

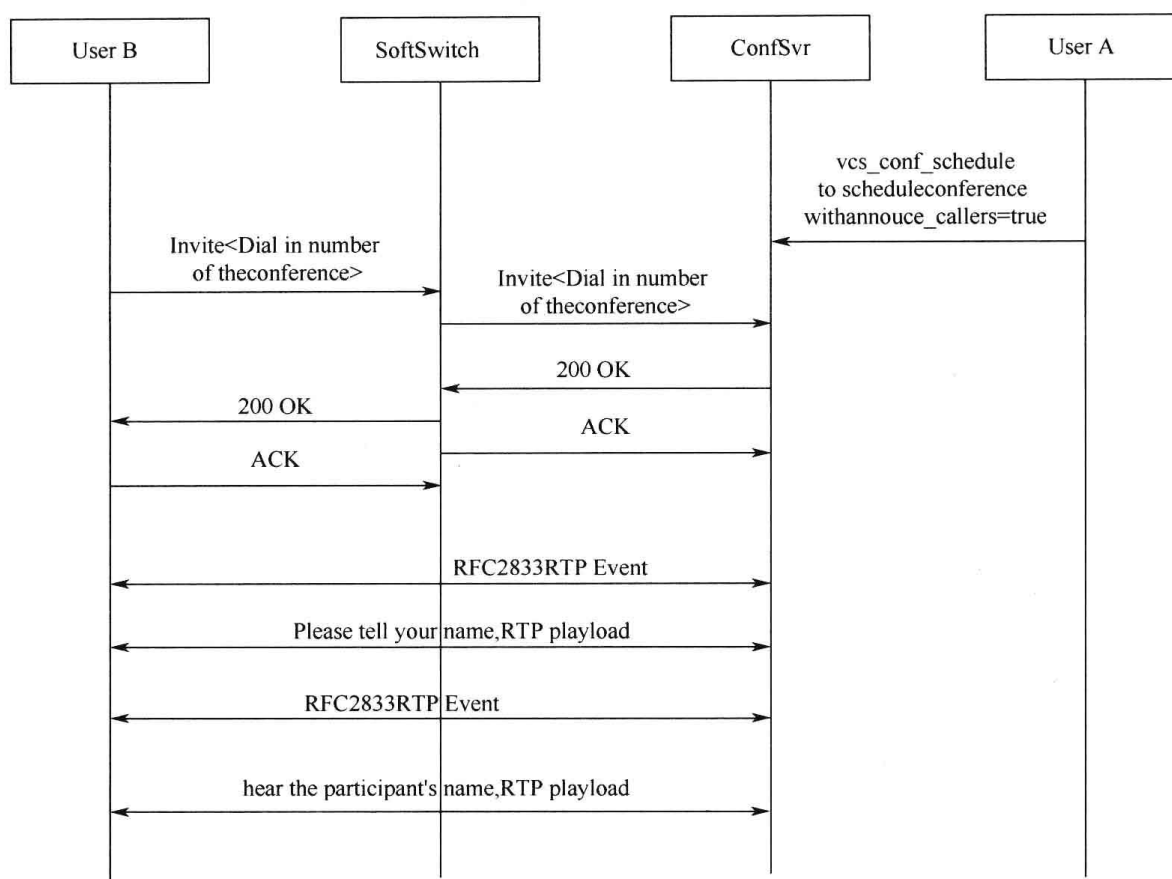


图 4.23 会议参加者通告

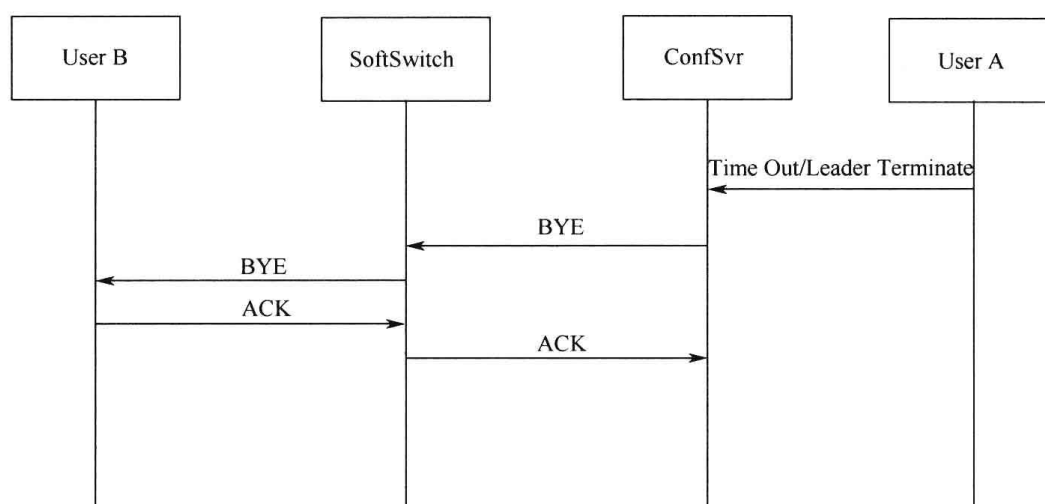


图 4.24 会议结束

(6) 数据协同

1) 应用共享

应用共享如图 4.25 所示，具体流程如下所述。

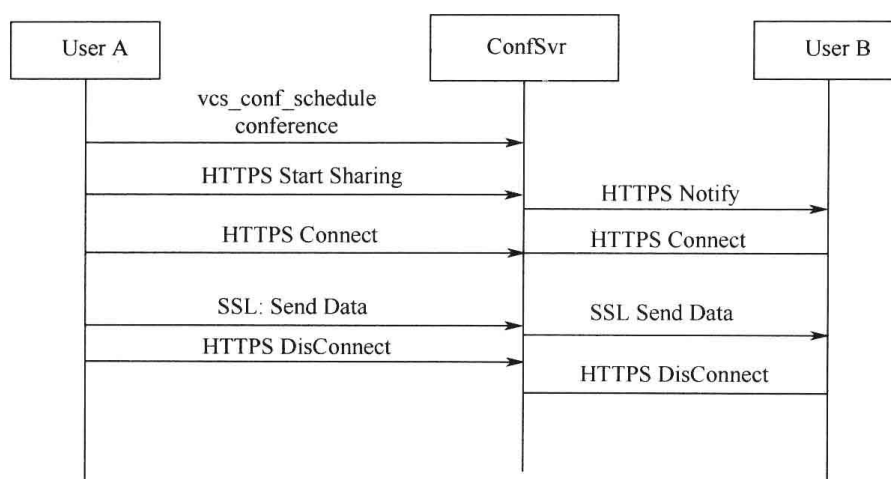


图 4.25 应用共享

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。
- User A 通过 HTTPS 通知会议服务器启动应用共享。
- User A 发送共享数据给会议服务器。
- 会议服务器及时通知 User B 客户端有共享数据。
- 会议服务器将共享数据推送到 User B 的客户端。

2) 演示模式

演示模式如图 4.26 所示，具体流程如下所述。

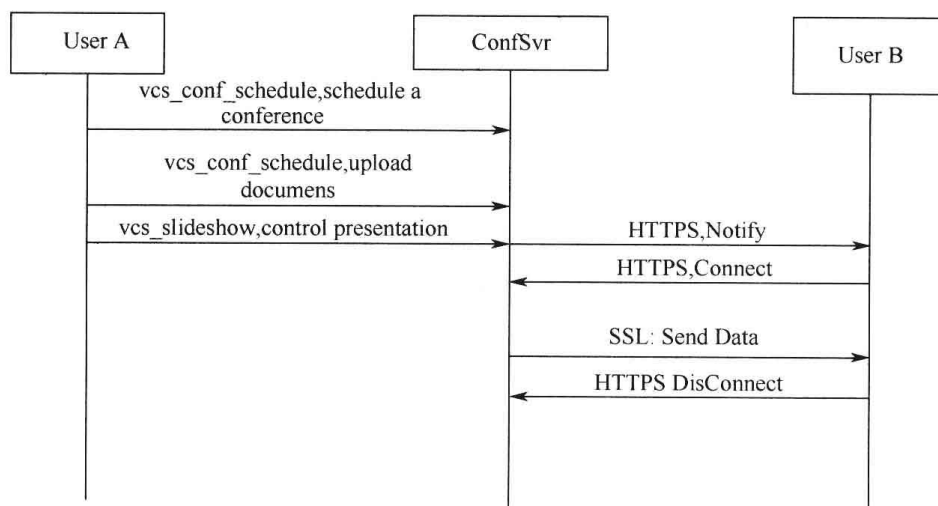


图 4.26 演示模式

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。User A 在预定会议时可以指定会议参数，如上传会议上需要演示的文件等。
- User A 可以通过 Web 对演示加以控制，如对 PPT 进行翻页等。

- 会议服务器会及时通知 User B 的客户端同步 User A 的演示内容。
- 会议服务器将演示数据推送给 User B 的客户端。

3) 桌面共享

参见应用共享。

4) 电子白板

参见应用共享。

5) 会议中即时消息

与应用共享的信令流程基本相同，请参照应用共享。即时消息并不依赖于会议，可以作为一个业务单独使用。

(7) 会议查询

用户直接在 Web 界面上可以查询会议历史记录。

(8) 会议回放

① 通过电话回放会议。

通过电话回放会议如图 4.27 所示，具体流程如下所述。

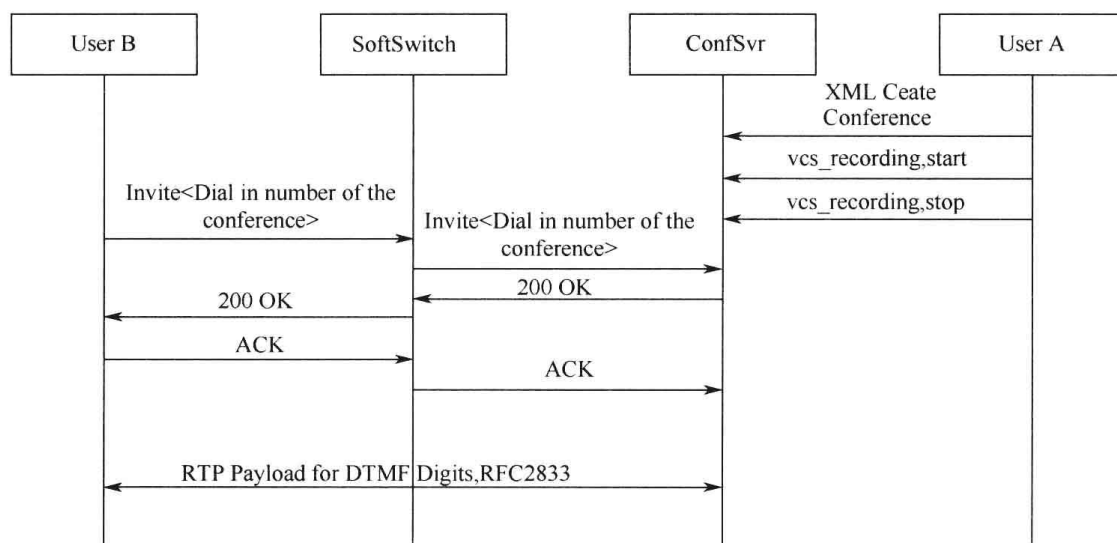


图 4.27 通过电话回放会议

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。
- User A 通过 Web 在会议服务器上录制会议内容并通知收听会议记录的拨入号和接入码。
- User B 通过会议终端拨收听会议记录的拨入号，通过软交换将收听会议记录请求传递到会议服务器，会议服务器将要求参与者拨入接入码。如果会议终端是 SIP 电话，那么会议终端通过 SIP INVITE 请求拨入收听会议记

录请求。

- User B 通过会议终端收听会议的接入码，即可收听会议记录。

② 通过 Web 回放会议。

通过 Web 回放会议如图 4.28 所示，具体流程如下所述。

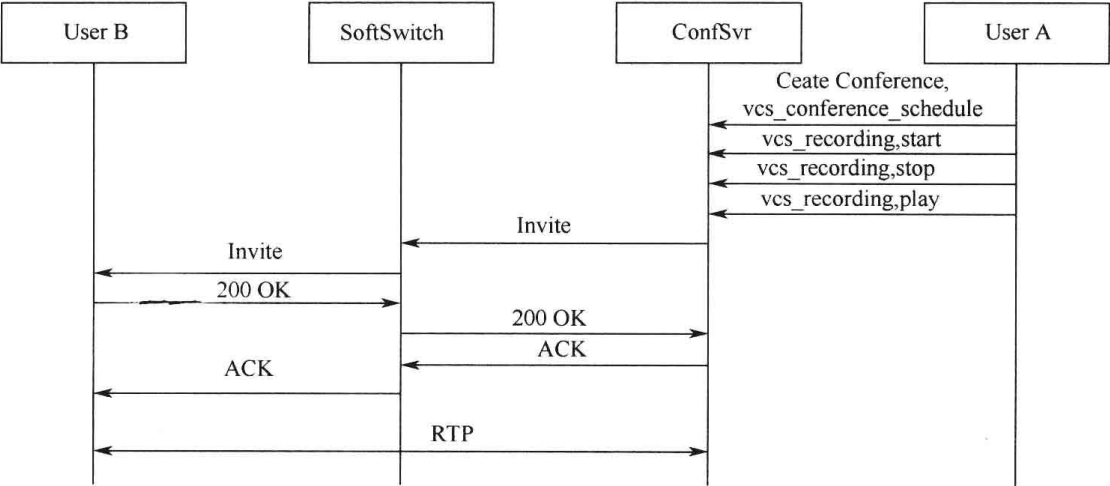


图 4.28 通过 Web 回放会议

- User A 通过 Web 在会议服务器上预定会议，预定会议时需要确定会议参数，如会议类型、会议主题、会议模式等。
- User A 通过 Web 在会议服务器上录制会议内容。
- User A 通过 Web 邀请 User B 收听会议记录。
- 会议服务器通过 SIP INVITE 请求触发软交换呼叫 User B 收听会议记录。
- User B 应答后收听会议记录。

4.3 基于 IMS 技术的会议

4.3.1 IMS 的概述

IMS 的英文全称是 IP Multimedias Subsystem，最初是 3GPP 组织制定的 3G 网络核心技术标准，目前已经被 ITU-T 和 ETSI 认可，纳入 NGN 的核心标准框架，并被认为是实现未来 FMC（固定 / 移动网络融合）的重要技术基础。IMS 是一个全 IP 的网络架构，采用 SIP 协议进行控制，实现移动性管理、多媒体会话信令和载体业务传输，实现端到端的 IP 业务。此外，IMS 还具有普通业务采用归属服务控制、接入独立性、业务控制、承载层完全分离等特点。

IMS 是一个开放的控制架构，目的就是在全 IP 网络环境中，为用户提供端到

端、可控的、有完全 QoS 保证的 IP 多媒体业务。

IMS 网络采取了分层架构进行体系设计。基于 IP 的传输和承载服务从 IMS 的信令网络和会话管理服务中分离出来，更高层的服务都运行在 IMS 信令网之上。分层的架构设计体现了“业务与控制分离”和“控制与接入或承载分离”的思想，减小了各层之间的依赖性，使得不同的用户终端能够通过不同的无线或者有线接入技术接入 IMS 网络，享受统一的呼叫控制服务及其相应的增值业务。这种“层次化”的网络构架设计为不同网络的互连互通和业务的融合奠定了基础。

IMS 的分层体系结构如图 4.29 所示。IMS 网络主要分成传送层、控制层、业务层三个层面。

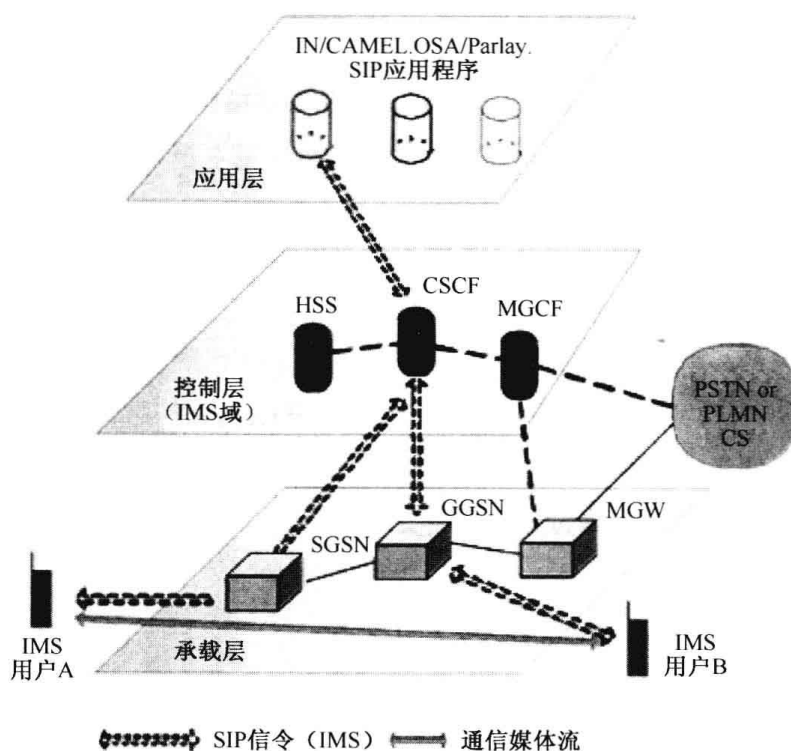


图 4.29 IMS 分层体系结构图

1. 传送层

传送层以 IPv4/IPv6 为核心技术，理论上支持任何 IMS 终端通过安装 IMS 会议客户端接入 IMS 多媒体会议系统，接入方式可以有 2G 移动（包括 GPRS、EDGE 和 cdma2000 1x）、3G 移动（包括 WCDMA、TD-SCDMA 和 cdma2000）、xDSL、WLAN 以及 LAN 接入等，但所有的媒体流均以 IP 分组的形式在网络中传输，网络运营商构建统一的 IP 宽带网络为所有的用户服务，体现以 IP 为核心的网络融合。

3GPP R5/R6 定义了移动接入，即通过服务 GPRS 支持节点（SGSN）和网关

GPRS 支持节点 (GGSN) 接入 IMS 网络: 固定接入, 如 DSL 接入则通过数字用户线接入复用器 (DSLAM) 和宽带接入服务器 (BRAS) 接入 IMS 网络; 基于 WiFi 的 IP 接入通过无线接入通路 (WAG) 接入 IMS 网络。网络附着子系统 (NASS) 和资源接纳控制子系统 (RACE) 完成对固定接入的接入控制。

2. 控制层

控制层以 SIP 为核心技术, 向下对传送层的 IP 多媒体通信进行呼叫控制和管理, 向上为业务应用层提供调用通信网能力的开发接口, 实现业务层和网络层的分离, 支持独立于网络运营商的业务运营商的形成, 体现了多媒体业务的统一控制与管理。

控制层的核心功能是呼叫会话控制功能 (CSCF), 负责 IMS 会话控制。CSCF 包括 P-CSCF (代理 CSCF)、I-CSCF (询问 CSCF)、S-CSCF (服务 CSCF) 等逻辑功能。P-CSCF 是 UE 接入 IMS 网络的入口点, 在 UE 和 P-CSCF 之间需要建立安全关联和 SIP 信令的压缩 (固定网接入不需要信令压缩)。S-CSCF 负责对 IMS 用户的鉴权, 执行会话控制业务并维护每一个会话的状态。I-CSCF 在 IMS 核心网中起到关口节点的作用, 提供域内 IMS 用户的 S-CSCF 分配、路由查询以及不同 IMS 域间的拓扑隐藏等功能。

控制层还包括突破网关控制功能 (BGCF)、媒体网关控制功能 (MGCF) 和信令网关 (SGW) 等用于与 PSTN 或 PLMN CS 域互通的功能。BGCF 完成到 PSTN/ISDN 的路由功能; MGCF 完成 SIP 与 H.248 的互通, 并通过 H.248 对媒体网关实现控制; SGW 完成 IP 与 CS 域信令传输层的互通。

媒体资源功能 (MRF) 提供语音和视频会议的必要支持, 包括控制部分和用户平面的处理部分 (如音频 / 视频流的混合)。IMS 与其他 IP 网络之间的互通功能由应用 (ALG) 和转换网关 (TrGW) 实现。ALG 工作在应用层, 为 SIP 消息进行 IPv6 和 IPv4 地址的翻译转换; TrGW 则在用户平面实现网络地址 (端口) 转换功能。

3. 业务层

业务层以当前信息技术为核心, 业务运营商可以通过部署在各类应用服务器中的业务逻辑向端用户提供增值业务, 业务提供独立于用户的接入技术, 体现了以信息与通信技术为核心的业务融合。

业务层的实体是应用服务器, 应用服务器是提供 IP 多媒体业务的实体。所谓 IP 多媒体业务, 是指基于 IETF 定义的会话控制能力, 利用分组交换域和多媒体承载来实现的网络功能。如图 4.30 所示, 说明了 IMS 中各种应用服务器的逻辑结构。

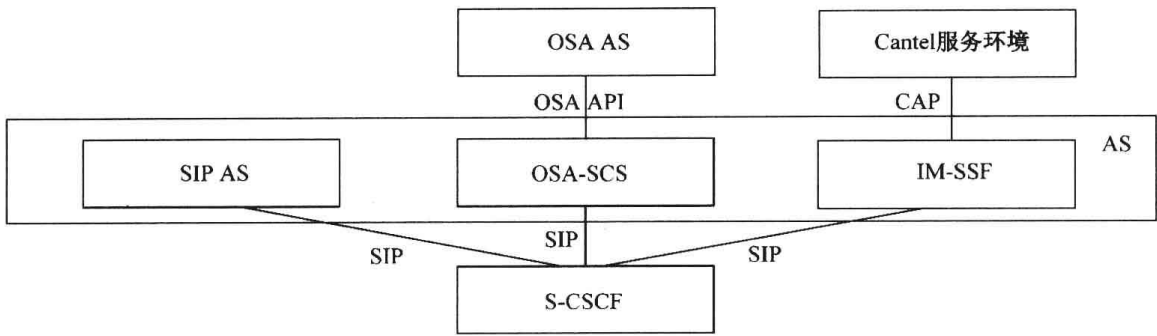


图 4.30 IMS 中的应用服务器

根据所支持业务类型的不同，应用服务器（AS）可以分为 SIP 应用服务器（SIP AS）、基于 Parlay 的 OSA 服务能力服务器（OSA SCS）和提供传统移动智能网业务功能的 CAMEL IP 多媒体服务交换功能（IM-SSF）。从 S-CSCF 的角度看，SIP AS、OSA 服务能力服务器和 IM-SSF 显示出相同的参考点行为。

SIP AS 是基于 SIP 的应用服务器，用于提供各种各样的增值业务。下文所介绍的会议就是基于 SIP AS 实现的。

4.3.2 IMS 会议结构

结合 IMS 的特点以及便于运营和集中控制的要求，在 IMS 架构上最合适的模型为紧耦合集中式服务器模型，通过会议服务器进行集中控制，每个会议可以通过会议服务器的一个会议中心来管理和控制在线的会话。

如图 4.31 所示，在本模式中，各成员间的通信通过核心的混合器来实现。每个成员只需与混合器建立媒体和信令的连接即可，由混合器对所有成员进行控制，并进行媒体流的混合分发。该模式适用于 Ad Hoc 型或预约好的拨入 / 拨出型会议。拨入型会议的用户可以通过访问会议控制节点动态加入会议，拨出型会议则需要通过 Web 等方式事先预定会议，由会议的中央控制节点邀请用户加入会议。Ad Hoc 方式则可以将正在进行的 SIP 对话自动迁移至会议状态，从而实现多方会谈。

在实际应用中，混合器相当于当前 H.323 会议中的 MCU 设备。本模式中任意一方退出会议，都不会影响其他成员会议的进程。只是当用户规模扩大或者地域分布广阔时，需要考虑多混合器的情形。

在 IMS 系统中，可以通过一些特定的功能实体相互配合实现多媒体会议功能，这些功能实体主要包括用户设备（UE）、媒体资源功能控制器（MRFC）、媒体资源功能处理器（MRFP）、会议应用服务器（Conference AS）等，如图 4.32 所示描述了 IMS 会议的架构。

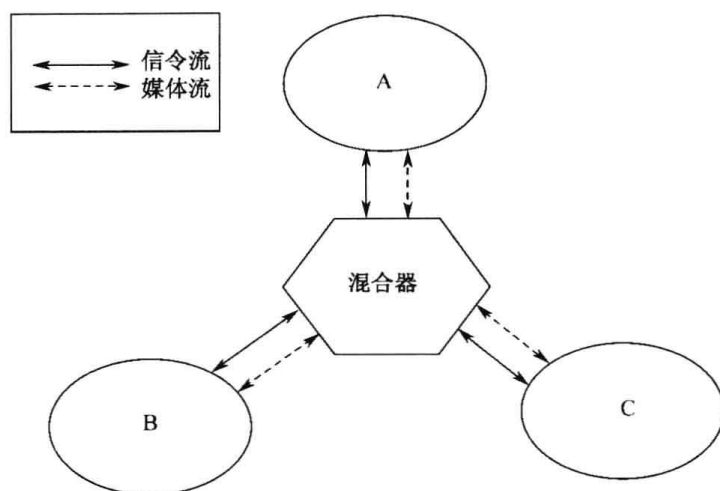


图 4.31 紧耦合集中式服务器模型

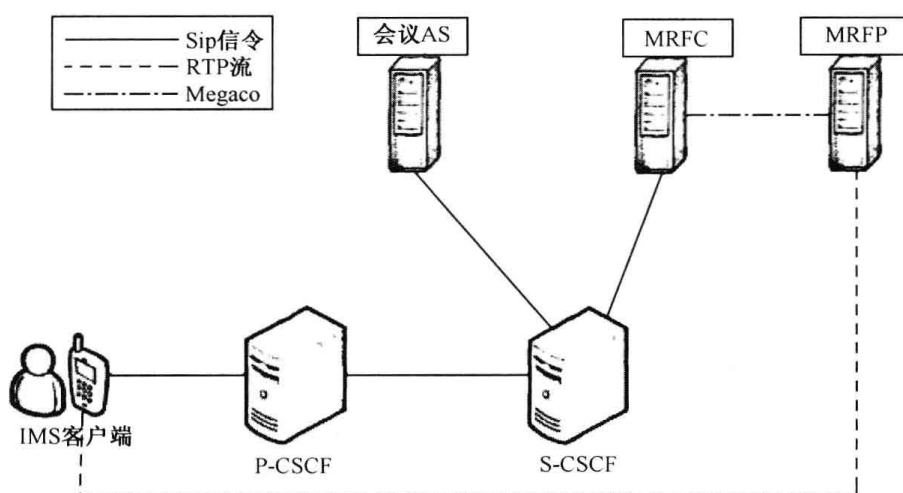


图 4.32 IMS 多媒体会议架构

（1）P-CSCF（代理 CSCF）

P-CSCF 是 IMS 终端盒 IMS 网络的第一个连接点，作为一个出站 / 入站 SIP 代理服务器，P-CSCF 将接入网中收到的 UA（User Agent）发来的 SIP 请求和响应沿着既定的方向继续传送（传向 IMS 终端或者传向 IMS 网络）。在 IMS 注册时，把一个 P-CSCF 分配给 IMS 终端，并且在注册过程中不会发生改变（IMS 始终通过一个 P-CSCF 进行通信），该节点也提供安全和 QoS 方面的保证。

（2）S-CSCF（服务 CSCF）

S-CSCF 是会话控制的中心节点，本质上是一个 SIP 服务器，但是它也会执行业务控制功能。除了 SIP 服务器功能之外，也提供注册、鉴权，这说明它保持用户位置（用户登录终端的 IP 地址）和用户的 SIP 地址记录（也被称为公共用户身份）之间的绑定。S-CSCF 总是位于归属网络，并作为中央控制节点根据网络运营

商的过滤规则提供服务。S-CSCF 提供了一个 ISC (IMS Service Control interface) 的统一的接口, 该接口采用了 SIP 协议。

(3) 会议 AS (会议应用服务器)

会议 AS 是本系统的核心功能, 是主持和执行会议服务的 SIP 实体。S-CSCF 根据过滤规则匹配, 如果匹配正确则转发 SIP 请求给该 AS, AS 根据相应的业务逻辑处理消息。

(4) MRF (媒体资源功能)

MRF 在归属网络中提供媒体信息源。MRF 支持归属网络播放媒体信息, 为多媒体会议提供混合媒体流, 进行编码转换和进行各种媒体分析。MRF 进一步可分为媒体节点 MRFP (媒体资源控制处理器) 和信令节点 MRFC (媒体资源功能控制器)。

- MRFP 执行所有与媒体相关的功能, 负责控制 RTP 媒体流的连接、混音、处理功能 (如音频的编解码转换)。
- MRFC 是 MRF 的控制实体。因为 IMS 会议是紧耦合会议并对媒体流采用集中控制, 所以可以将 MRFC 和 AS 部署在一起, MRFC/AS 通过 H.248 接口控制 MRFP 上的资源。

4.3.3 控制信令 SIP 概述

IMS 是以 SIP 为核心的网络, 因此, 深入理解 SIP 协议对于设计基于 IMS 的会议业务流程至关重要。SIP 协议详见前文。

4.3.4 媒体信令 H.248 概述

H.248 协议, 又称为 Megaco 协议, 应用于媒体网关和媒体网关控制器之间, 处理多媒体会议期间的信令和会话管理。媒体网关控制器和媒体网关构成主从关系。

1. 连接模型

H.248 协议的连接模型将媒体网关内的主要对象描述成终节点 (Termination) 和上下文 (Context), 它们可以由媒体网关控制器来控制。一个终节点发起 (Sources) 或终止 (Sink) 一个或多个媒体流, 并且各终节点都拥有实际媒体流的信息。

不同终节点之间通过上下文联到一起。不与其他终节点关联的终节点集合定

义为由特殊类型上下文（即 Null 上下文）表示的终节点。上下文描述与它关联的终节点的拓扑，例如，在上下文含有两个以上终节点的情形下，它包含与媒体混合相关的参数。

2. 协议操作

本协议含有一套用来操纵逻辑实体的命令，这些实体在连接模型中有描述。具体说来，H.248 提供的命令集有使用二进制格式或者文本格式两种方式编码。H.248 协议中定义了如下 8 条命令，命令控制对象都是终端。

① **Add**: 向上下文增加一个终节点，它还用来隐形地创建上下文（只有加入第一终节点，就创建一个上下文）。

② **Modify**: 修改终节点的状态特性和媒体流特有的特性。

③ **Subtract**: 从上下文去掉一个终节点，它还用来隐形地删除上下文。和创建终节点相似，去掉上下文的最后一个终节点将导致删除该上下文。

④ **Move**: 将某个终节点从一个上下文移到另一个上下文。

⑤ **AuditValue**: 返回当前终节点特性以及该终节点的事件、信号和统计。

⑥ **AuditCapabilities**: 返回某终节点特性的所有可能值以及某个媒体网关所允许的信号和事件。

⑦ **Notify**: 当某个事件发生时，由媒体网关用来向媒体网关控制器进行报告。

⑧ **ServiceChange**: 由媒体网关用来向媒体网关控制器报告有关服务的确切变化（如当某个特定终节点脱离服务或返回服务时）。

每个命令可以携带若干参数，称为描述符。命令也可以将描述符作为输出返回。

4.3.5 IMS 的会议系统特点

IMS 充分考虑了实际运营的需求，在 QoS、安全、计费等方面都制订了相关规范。IMS 架构是基于逻辑功能进行描述的，这给具体的物理实现以及逻辑功能组合带来了极大的灵活性。

基于 IMS 固定移动网络的融合架构的会议系统主要有以下特点。

（1）统一的业务和应用平台

统一的业务和应用平台使得所有的应用层资源得以在所有业务之间共享，如用户业务数据、在线状态、位置信息、好友列表等。

(2) 统一的业务控制

IMS 具有集中业务控制的特点, 因此, IMS 多媒体会议业务可以和其他多媒体业务一起在一个统一的平台上实现业务控制。架构于 IMS 系统之上的多媒体会议系统可以充分利用 IMS 统一用户鉴权、会话控制、QoS 控制、计费和安全等功能和机制, 极大地方便了运营商对多媒体会议业务的控制和管理。

(3) 统一的用户数据

利用 IMS 用户数据的统一管理功能, 架构于 IMS 上的多媒体会议系统可以实现业务的大容量、大范围推广和部署。传统模拟视频会议系统和基于 H.323、SIP 的会议系统具有较低的业务使用灵活度, 用户数据的管理也较为不灵活。特别是传统的专用视频会议系统, 只能够在特定的会议室用特定的终端设备使用会议业务, 而基于 IMS 的多媒体会议系统, 对于不同的接入方式, 都可以实现统一的用户数据管理, 使得同一个用户的账户, 可以在不同的地域通过不同的接入方式进入多媒体会议系统, 极大地增强了业务的使用灵活度。

(4) 更多的接入方式, 快速的业务部署

由于 IMS 是一种与接入无关的网络, 因此架构于 IMS 系统之上的多媒体会议系统可以最大限度地支持各种不同类型的接入方式, 扩大了多媒体会议系统的使用场合, 同时由于 IMS 实现了控制与承载、控制业务分离的分层架构, 使得 IMS 多媒体会议系统可以在 IMS 上实现快速部署, 并且降低了业务的部署成本。

(5) 更丰富的会议功能

相对于以前的基于 H.323 和 SIP 的会议系统, IMS 会议系统可以支持更多的媒体类型, 不但包括了常用的音视频会议, 也包括了如文字聊天共能, 使业务的应用场景进一步扩大。

4.3.6 会议系统需求分析

会议系统在 IMS 中可以提供多个呼叫连接的能力, 即在三个或者更多个 IMS 用户之间同时进行通话。任何一个非主控用户挂机, 其余用户照样保持原来的通话连接。

基于 IMS 的会议系统, 是在统一 IP 多媒体子系统架构上实现的会议业务, 通信媒体包括了音频、视频、消息及各种数据应用等, 例如, 实现文件和白板共享, 通过视频来传送相图像信息。

目前的会议系统有如下的分类。

(1) 按照会议系统是否有预约分类

可分为 Ad Hoc 会议和预约型会议。

- “Ad Hoc 会议”无须预先设定，在任何时间创建会议并立即召开一个会议。适合于一些临时的、非正式的会议。
- “预约型会议”需要事先进行预约会议召开时间，设定会议时长等策略，并对会议成员进行通知。由于采用预约的方式，保证了开会所需的系统资源可以被预占，从而保证会议的可靠性。适用于一些比较重要、比较正式的会议。

(2) 按照会议的密级分类

可分为普通会议和特殊会议。

- 普通会议：密级为普通的会议，一般不需要密码都可以进入。
- 特殊会议：一般是指只能限制几个级别较高的人员，涉及的会议内容密级较高，一般需要密码验证才能进入会议。

4.3.7 会议系统的功能

会议系统能够完成以下基本功能。

(1) 会议的预定、修改、取消等管理功能

- 会议的预定、修改、取消等管理功能，可通过 Web 界面对会议策略进行制订和修改，通过 Web 界面预定、修改取消会议。
- 会议的预约是指事先预约会议召开时间，设定会议时长，设定会议主席、参加会议的人员等会议策略，并通知会议成员。对预约会议，可以延长会议时间。

(2) 会议的创建和召开

会议系统支持两种形式的会议：预约型会议和 Ad Hoc 会议。因此会议创建方式即有两种：一种是预约会议资源，到了时间即自动创建会议；一种是由会议发起者即时地主动创建会议。

会议一旦创建将分配和预留相应的会议资源并产生该会议服务器上唯一的会议标识号。

会议支持多种方式启动会议：

- 参与者收到 E-mail 或短信通知后通过拨特服号和会议标识号进入会议，即用

户通过访问会议控制节点动态地加入会议（Call-in 型会议）；

- 需要通过 Web 的方式事先预定会议，当会议的召开时间到，由会议的中央控制节点主动邀请用户加入会议（Call-out 型会议）；
- Ad Hoc 会议方式则可以将正在进行的会话自动迁移至会议状态，从而实现多方会谈。

（3）会议的结束和删除

- 会议删除：与此会议相关的会议策略等资源应销毁，与会者都退出会议。
- 如果会议结束时会议中还有用户未断开连接，会议服务器应向这些用户发送信令，让他们全部断开连接后再将会议删除。

系统应该在以下三种情况下实施会议结束操作：

- 主席退出会议；
- 会议结束时间到；
- 对“即开会议”来说，所有用户都已经退出会议。

（4）Ad Hoc 会议控制

- Ad Hoc 会议控制主要完成：查询所有和会议及参与人相关的信息和状态，编辑，添加 / 删除参与人等功能。
- 查询会议状态：对于订阅会议状态信息的用户，当会议状态改变或者有新入参加会议等情况，会通过短信通知各个订阅用户。
- 添加和删除参与者：可以是自己加入或退出一个会议，也可以由第三方（可以是主席，调度员和其他与会者）控制将自己加入或退出会议。

4.3.8 会议系统在 IMS 中的位置

会议系统在 IMS 系统中是作为一种 SIP AS 提供的，位置如图 4.33 所示。

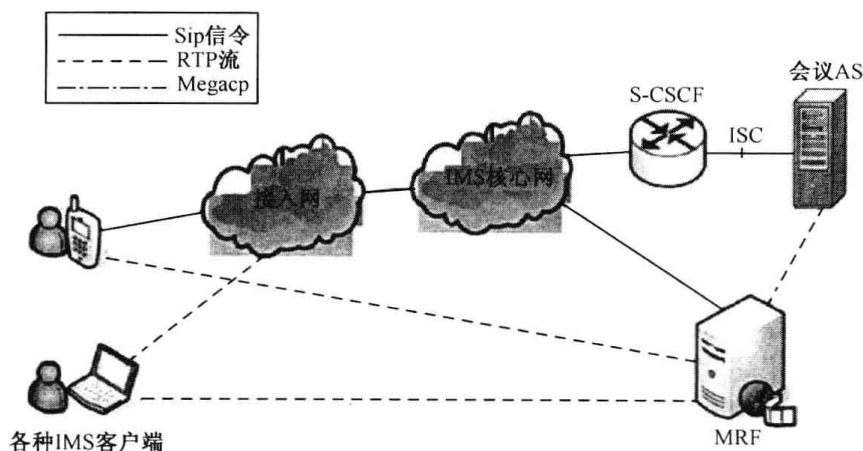


图 4.33 会议系统在 IMS 中的位置

会议 AS 模块需要和 S-CSCF 模块通过 ISC 接口完成信令交互，以使得会议业务能够通过 IMS 提供给终端。会议 AS 对 IMS 而言，相当于一个业务终端，只是在这样一个物理实体上可能有许多终端，如会议服务器，里面有许多会场（服务终端）。

会议 AS 模块通过 H.248 命令依据一定的媒体策略控制媒体服务器完成媒体的处理和分布。媒体服务器采用纯软件的方式实现媒体混合以及转发等操作。

用户通过登录会议 AS 中提供的 Web 服务来完成会议预约、发送会议通知等功能。

1. 会议系统功能模块

会议系统划分为以下几个功能模块，如图 4.34 所示。

- 会议中心；
- 会议通知服务模块；
- 会议和媒体策略服务器；
- 数据应用服务器。

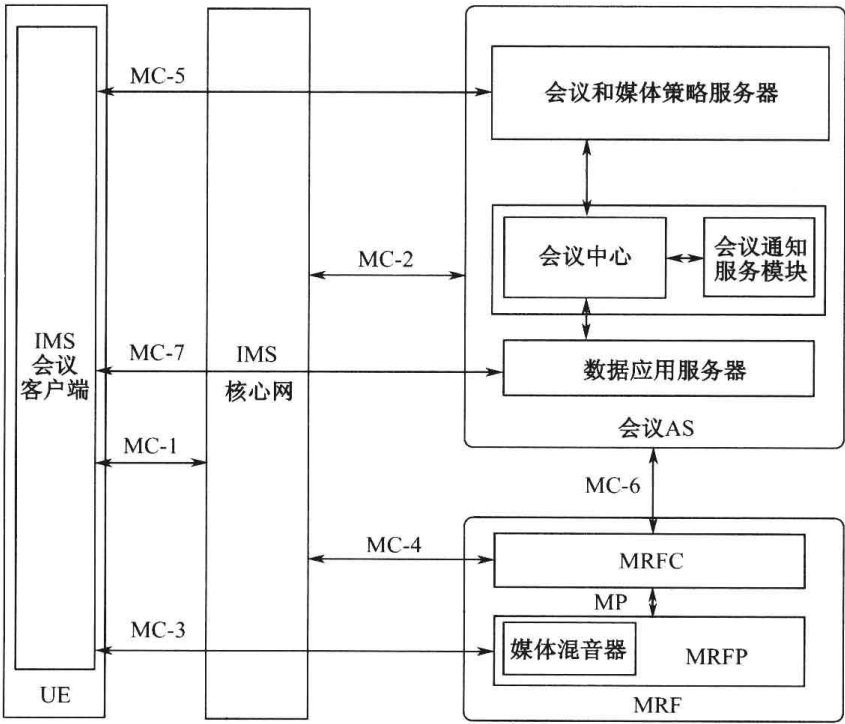


图 4.34 会议系统架构

(1) 会议中心

它是整个会议系统的核心模块，实现 Focus 功能，主要完成信令的处理，以及对会议系统中的其他各个模块的控制。一个会议只有一个会议中心，会议中心由

会议 URI 唯一标识，具体功能如下所述。

- 会议的预约处理：从会议管理模块获取预约信息并且进行会议的预约工作；
- 启动注册会议功能：负责会议的启动和结束工作；
- 会议策略的执行功能；
- 根据会议媒体策略，控制整个会议媒体流混合和分发；
- 会议状态的控制以及会议状态的通告；
- 会议的资源管理。

(2) Ad Hoc 会议控制

① 会议通知服务模块。

会议通知服务模块的工作机制由 RFC3265 定义。会议通告服务主要完成向与会人员通知相应的会议事件和会议状态。与会者可以通过发送 SUBSCRIBE 请求向会议通知服务器预定感兴趣的会议事件和状态通知服务。在会议过程中，一旦预定的事件发生，会议通知服务器将通过 NOTIFY 方法向与会者通告相关会议事件和当前会议状态。例如，用户可以预定会议在席信息，当有新的与会者加入或有人退出会议时，会议通知服务器将向预定会议在席信息的用户发送会议成员变更情况。会议状态信息可以用 SML 文档描述。

② 会议和媒体策略服务器。

会议和媒体策略服务器主要用于生成、存储和操作会议策略。一般而言，一个会议对应一个会议策略（包括会议成员策略、会议策略，媒体策略），是客户端和会议 / 媒体策略之间的接口。

会议和媒体策略服务器也可以接受订阅，当存储的策略 XML 文档发生变化时，会议 / 媒体策略服务器会将相应信息通知订阅用户。

会议中心是会议策略的具体执行者，会议中心使用会议策略控制和管理整个会议。

- 会议策略：存储会议策略信息以及相关会议参数信息，包含会议类型是预约型还是 Ad Hoc 型、会议中所能容纳的最多的人数、会议进入是否需要密码、会议的开始时间、结束时间等。
- 会议成员策略信息：会议成员策略主要用于会议成员的管理，包括用户接入控制、用户权限管理、用户标识、用户认证密码、成员权限等属性，以及会议成员在加入过程中是否要鉴权等。
- 媒体策略：包括媒体混合策略、同步策略、媒体分发策略、用户标识以及音频的编解码信息，会议控制模块通过媒体策略信息控制媒体混合器进行该用户的媒体处理和分发。

③ 数据应用服务器。

数据应用服务器是实现会议中数据应用相关功能的实体，实现的功能包括电子白板共享、文字聊天等数据协同功能。

2. 外部接口与协议

(1) MC-1 接口 MC-1 接口

该接口支持会议客户端到 IMS 核心网之间的通信，基于 SIP 协议，满足 IMS Gm 接口要求，主要支持如下功能：

- 会议客户端和 IMS 核心网络之间的 IMS 会议的会话信令；
- 执行会议用户的鉴权和授权；
- 提供会议客户端的注册。

(2) MC-2 接口

该接口支持 IMS 核心网和会议应用服务器之间的通信，实现会话控制；该接口基于 SIP 协议，满足 IMS ISC 接口的要求，主要支持如下的功能。

- 提供会议服务器和 IMS 核心网络之间的 IMS 会议的会话信令；
- 会议服务在 IMS 中设置参数。

(3) MC-3 接口

该接口支持会议客户端与 MRFP 之间的媒体传输、资源接入控制功能。实时媒体的传输使用 RTP/RTCP 协议，资源接入控制协议采用 BFCP 协议。MC-3 具体支持如下功能：

- 媒体传输和发送控制过程；
- 媒体传送的状态报告；
- 资源接入控制消息的传输。

(4) MC-4 接口

该接口支持 IMS 核心网到 MRFC 之间的通信，使用 SIP 信令，除满足 IMS Mr 接口要求外，主要支持向 MRFC 中转来自会议服务器的媒体控制信息。

(5) MC-5 接口

该接口为会议客户端到会议服务器之间的接口，采用 HTTP 协议，主要对用户会议 XML 文档的管理，包括创建、修改、获取和删除等。

(6) MC-6 接口

该接口是会议服务器到 MRFC 之间的接口，主要用于会议服务器对媒体资源

的控制；该接口基于 SIP 协议。

(7) MC-7 接口

该接口为会议客户端到数据应用服务器的接口，主要为客户端提供会议数据相关功能；该接口基于 HTTP 协议。

(8) Mp 接口

Mp 接口为 MRFC 到 MPRP 之间的接口，基于 H.248 协议，对该接口的详细描述参考 3GPP TS23.333 和 TS29.333，主要功能包括：

- MRFC 对 MRFP 中多媒体会议资源的控制，承载资源接入控制请求或响应消息；
- 媒体混合的控制、传递来自会议中心到媒体混合器的控制命令，传送混合器的反馈给会议中心等。

4.3.9 信令流程

1. 创建会议

用户可以通过 Web 方式预约会议，根据用户需求，由会议中心预留会场资源并配置会议参数（会议类型、参加人员、起止时间、规模、媒体类型等），创建的会议由会议中心分配的 URI 来唯一标识，并返回给会议创建者，或者按照一定策略通知会议成员（E-mail、公告等）。用户在会议启动前使用该唯一的 conf-URI 加入。用户也可以使用它们自定义的“conf-URI”请求创建会议，如果请求的 URI 没有语法错误，也不会引起冲突，会议于媒体策略服务器可以接受此 URI 为新创建会议的 URI，即 Ad Hoc 方式。

创建会议的信令流程如图 4.35 所示。

① 会议客户端向 IMS 系统发送 INVITE 请求，请求创建一个 Ad Hoc 会议，该请求被转发到用户归属网络的 S-CSCF，S-CSCF 通过与 HSS 交互，检查用户的业务配置和触发条件，在满足业务触发条件的情况下，将业务触发到相应的会议业务服务器，如图 4.35 中①～⑤步骤所示。

② S-CSCF 将 INVITE 请求转发相应的会议业务服务器，会议服务器首先判断正在进行的 Ad Hoc 会议的数目，若小于系统上限则按照系统默认的会议参数生成一个新的会议 Focus（包括与其对应的 URI），并向注册模块登记新生成的会议 URI，同时向 MRF 发起一个 H.248/SIP 交互，为该用户创建一个 IMS 连接点，确定 MRFP 的媒体资源能力，如图 4.35 中⑥～⑨步骤所示。

③ 会议服务器向客户端返回一个 183 临时响应，如图 4.35 中⑩～⑬步骤所示。

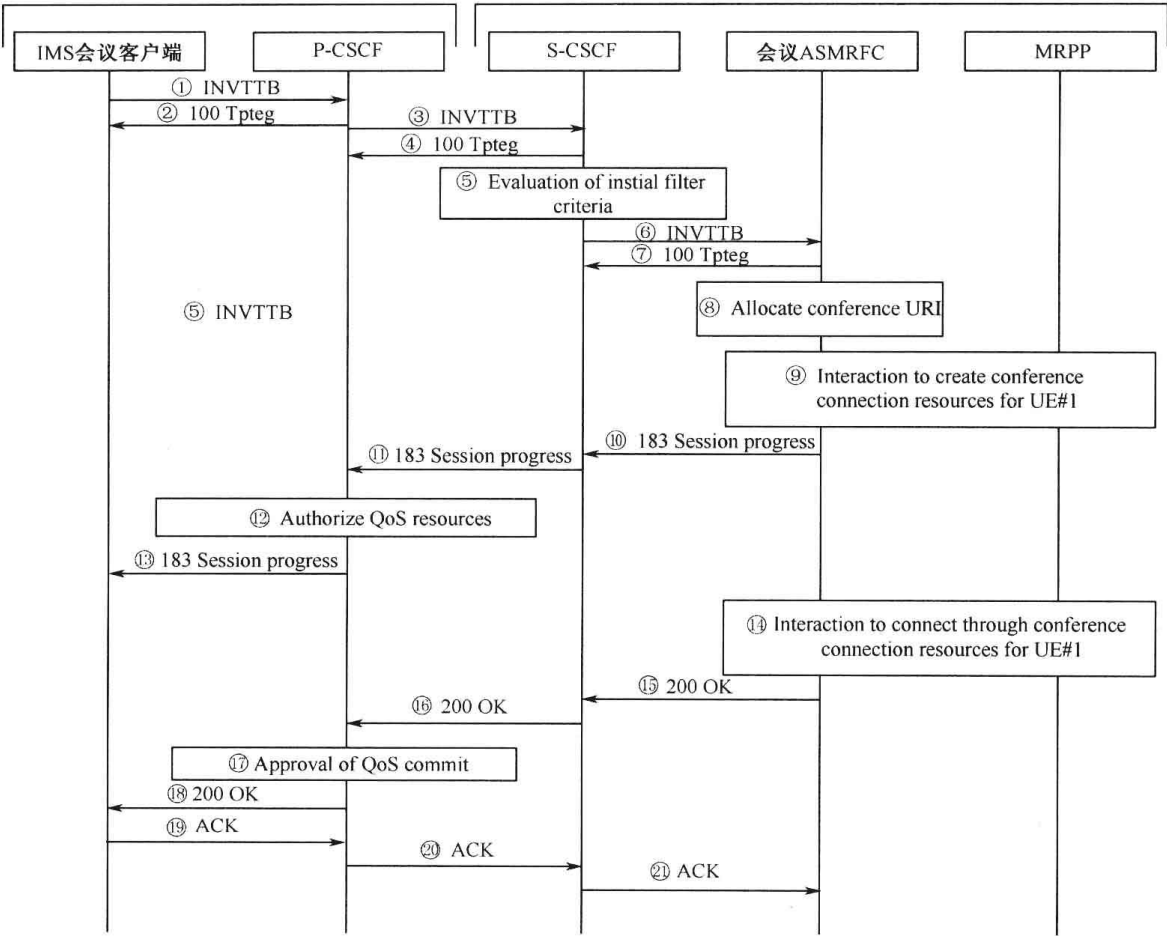


图 4.35 创建会议的信令流程

④ 根据会议客户端和服务器的协商结果，服务器发起一个 H.248/SIP 交互，并指示 MRFP 为用户进行多媒体处理资源预留，如图 4.35 中⑮～⑲步骤所示。当资源预留结束时，会议服务器向客户端返回 200 OK 确认响应。响应中包含了为该会议分配的 URI 等参数，同时会议服务器发起一个 H.248/SIP 交互修改连接，将会议客户端连接到 MRFP 中预留的多媒体处理资源，如图 4.35 中⑭～⑮步骤所示。

⑤ 客户端启动媒体流，并向会议服务器发送 ACK 请求。

2. 成员加入

成员加入可以是第一方的，也可以是第三方的。前者指用户自己加入，后者是指一个用户增加另外一个用户。用户可以通过多种方法获得会议的 URI，例如，通过即时消息、E-mail、短信等。

(1) 用户主动加入

用户可以向会议 URI 发送 INVITE 请求，如果会议策略允许其加入，就被加入到会议中。

如图 4.36 所示，会议客户端 UE 向会议 AS 发送 INVITE 消息请求加入会议，会议 AS 在 200 OK 响应中的 Contact 头域中包含了会议 URI 和 “isfocus” 参数，用于标识加入的是一个，“isfocus” 参数指示这是一个焦点 URI。这个会议 URI 可以用来指引其他用户到会议中，也用于订阅会议状态事件包，并在 SDP 中进行媒体协商。

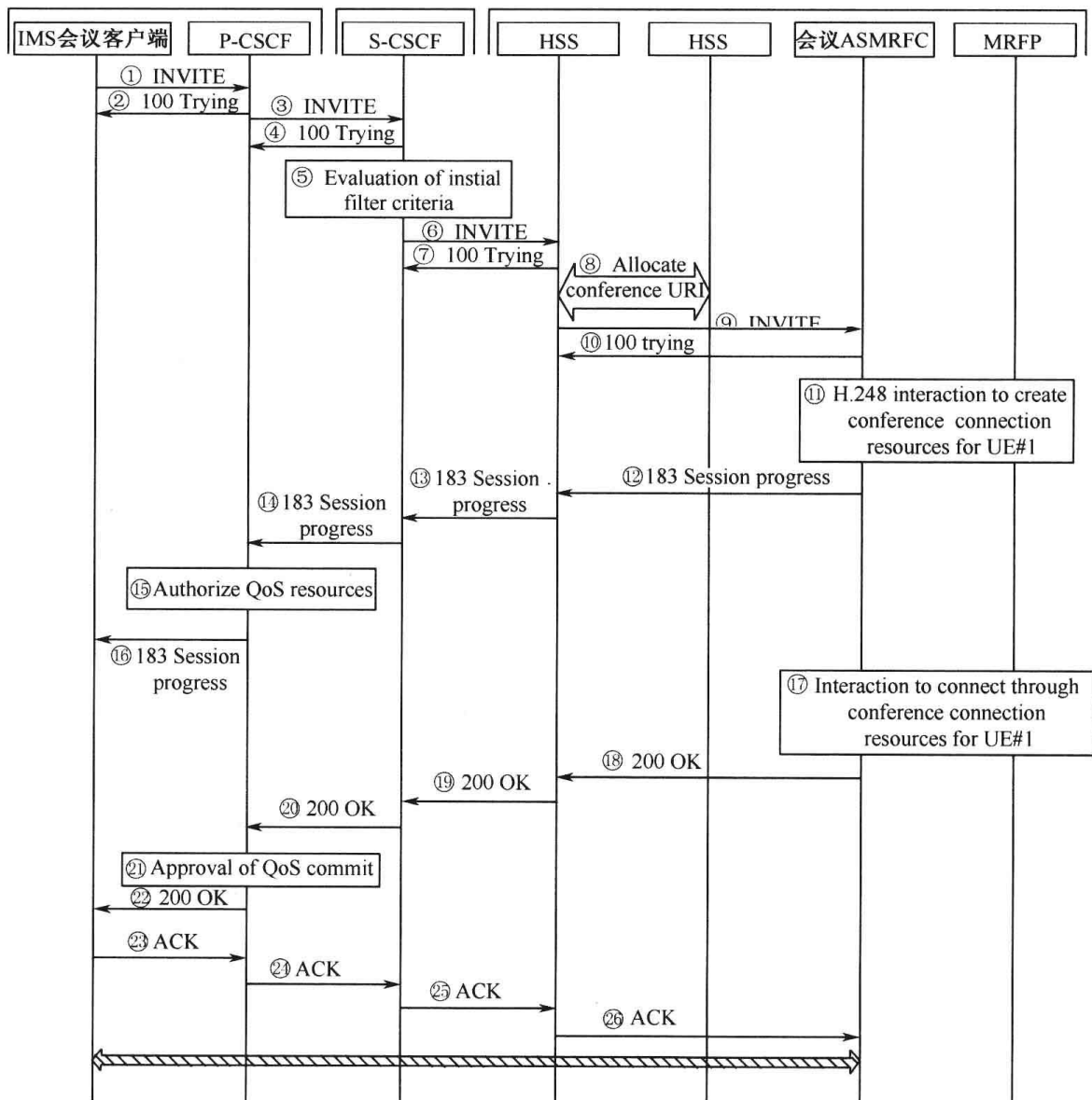


图 4.36 用户使用 Conference URI 加入会议

(2) 用户受邀加入

利用 SIP 方法实现第三方加入，可以使用 REFER 消息。有两种情况，一种由会议成员向要加入的用户发送 REFER 消息，告诉他向会议 URI 发送 INVITE 消息加入会议；另一种情况是会议成员向会议 AS 发送 REFER 消息，要求其向要加入的用户发送 INVITE 消息（Dial Out）。后一种情况允许成员把不支持 REFER 消息

的会议不敏感用户加入到会议中来,如图 4.37 所示。

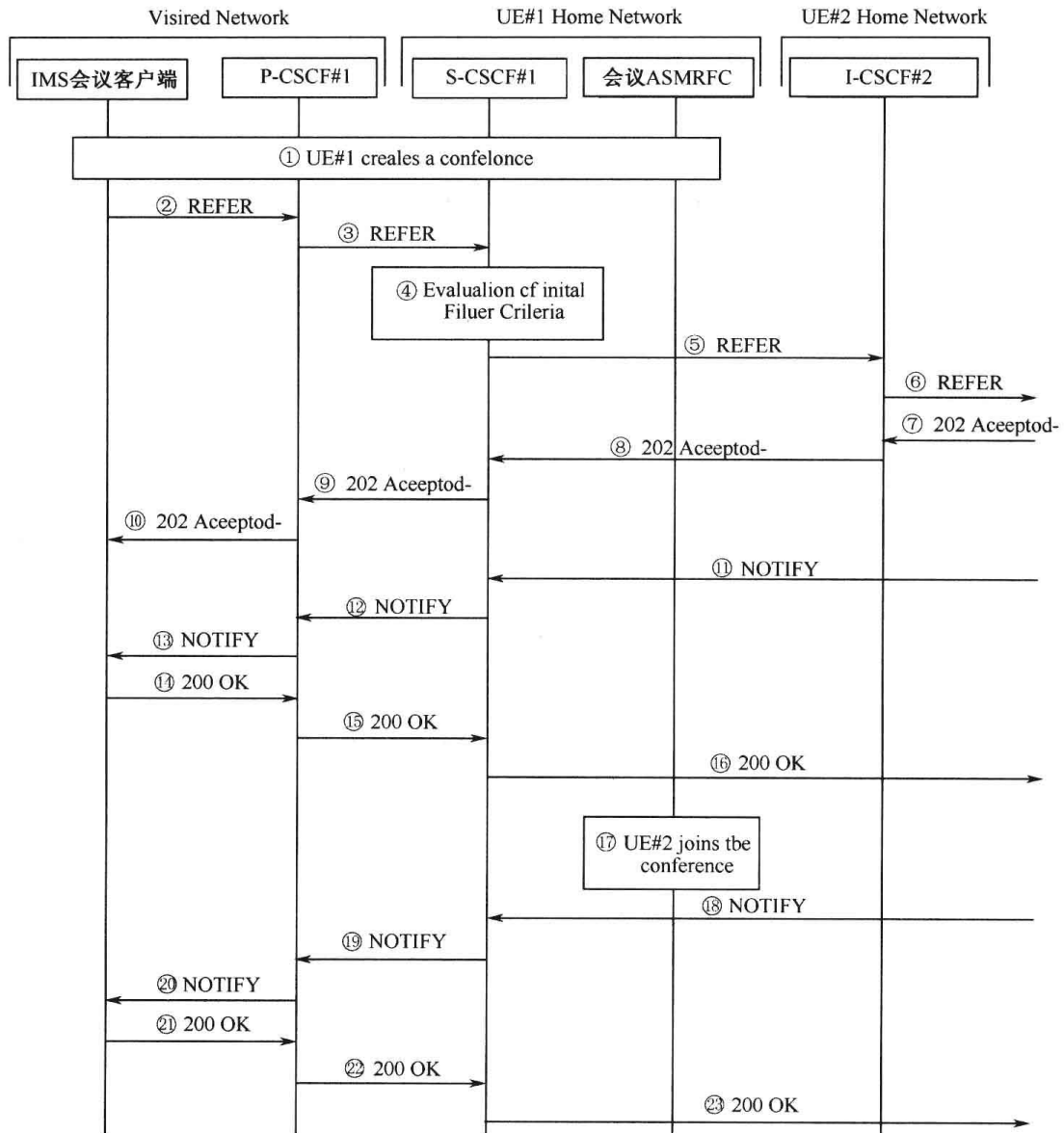


图 4.37 用 REFER 邀请用户加入会议

IMS 会议客户端 UE1 创建会议成功后,得到会议的 URI,发送 REFER 消息给 UE#2,指示其加入会议。REFER 请求的 Refer-TO 消息头设置为会议的 URI,并包含一个“isfocus”参数。S-CSCF 分析目标地址并决定转发的目标网络。I-CSCF 向 HSS 执行位置查询找到 UW2 的 S-CSCF。UE2 的归属网络在收到 REFER 请求后发送 202 (Accepted) 响应并开始向 UE2 逐级转发该请求。UE2 收到该请求后按照 RFC3515 来处理,并向 UE#1 通报 (NOTIFY) 呼叫进程。

(3) CPCP 方法

CPCP 是一种被用户用来控制会议策略的客户端 / 服务器协议。会议策略由一

个 URI 标识，每个会议都有一个唯一的会议策略 URI，该 URI 指向一个会议策略服务器（CPS），在此用户可以控制会议策略。必要时，可将会议策略的变更通知会议 Focus，从而完成响应的操作。例如，用户在会议的 DL 种增加新的用户地址，CPS 将此通知 Focus 后，Focus 会发送 INVITE 消息邀请该用户加入会议。

CPCP 的一个基本功能是向会议中增加成员，这相当于修改成员列表，属于会议控制范畴。CPCP 客户可以向会议 URI 发送请求，指定要加入用户的 URI（可以是其自身的 URI），如果此 URI 所标识的用户不在会议中，会议控制模块就向这些用户发送 INVITE 消息，使之加入会议，其信令过程不再详述。CPCP 方法如图 4.38 所示。

```
<xs:complexType name="floor Type">
  <xs:sequence>
    <xs:element name="resources" type="resourcesType" minOccurs="0"/>
    <xs:element name="urces" type="urcesType" minOccurs="0"/>
    <xs:element name="moderators" type="moderatorsType"
      minOccurs="0"/>
    <xs:arly namespace="##pther" processContents="lax" minOccurs="0"
      maxOccurs="unbounded"/>
  </xs:sequence>
  <xs:attribute name="maxHolders" type="xs:int" default="1"/>
</xs:complexType>
```

图 4.38 CPCP 方法

上述 floorType 的 XML schema，它包含三个子元素。“resources”的类型为“resourcesType”，它列出了这个 floor 所包含的资源。如果在 floorType 中未列出这一项，表示该 floor 包含该会议的所有资源。“users”的类型为“usersType”，它定义了能够使用该 floor 的使用者。如果该项未列出，则表示该会议的所有用户都能够使用该 floor。“moderators”的类型为“moderatorsType”，它定义了对该资源有控制权的主席。如果在“CreateFloor”命令中没有设定，则表示发送“CreateFloor”命令的用户是控制该资源的主席。

3. 订阅会议信息

3GPP TS 24.147 中定义了 SUBSCRIBE/NOTIFY 消息，会议成员可以向会议 URI 发送 SUBSCRIBE 消息订阅会议通告服务，接收包含成员列表等消息的通告。当有新成员加入或者有成员离开会议时，会议服务器会发送 NOTIFY 消息通告订阅者会议状态的改变情况。

用户订阅会议信息如图 4.39 所示。

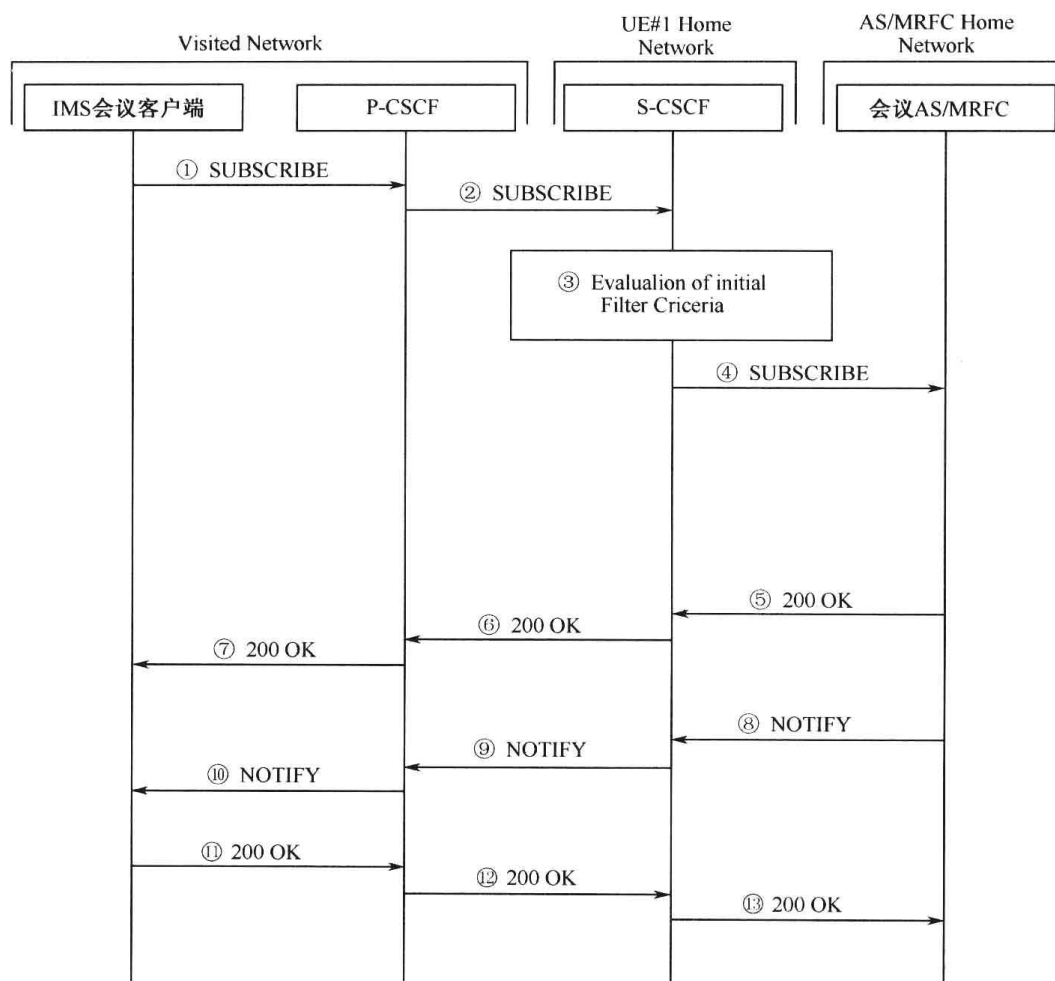


图 4.39 用户订阅会议信息

4. 成员离席

成员离席和成员加入有相同之处，也分为第一方和第三方离席。第一方离席可以通过会议 URI 发送一个 BYE 请求完成，第三方离席可以通过 REFER 方法完成。在这种情况下，已授权的 SIP UA 可以向会议 URI 发送 REFER 请求，其中的 Refer-To 头域包含了要离席的成员的 URI，并将其中的方法设定为 BYE。发送请求者无须知道要离席的成员和会议之间的对话消息，会议指导此信息，并在生成 BYE 消息时将相应信息填进去，如图 4.40 所示。

另外，对应对于成员加入，也可以通过管理员授权用户使用 CPCP 方法来删除任何成员（包括他自身）。客户可以使用成员管理操作命令，向会议发送请求，指定要删除的成员 URI，触发其向要删除的成员的 URI 发送 BYE 请求。CPCP 方法可以用来删除大量用户，一般是指“群删”（Mass Ejection）。

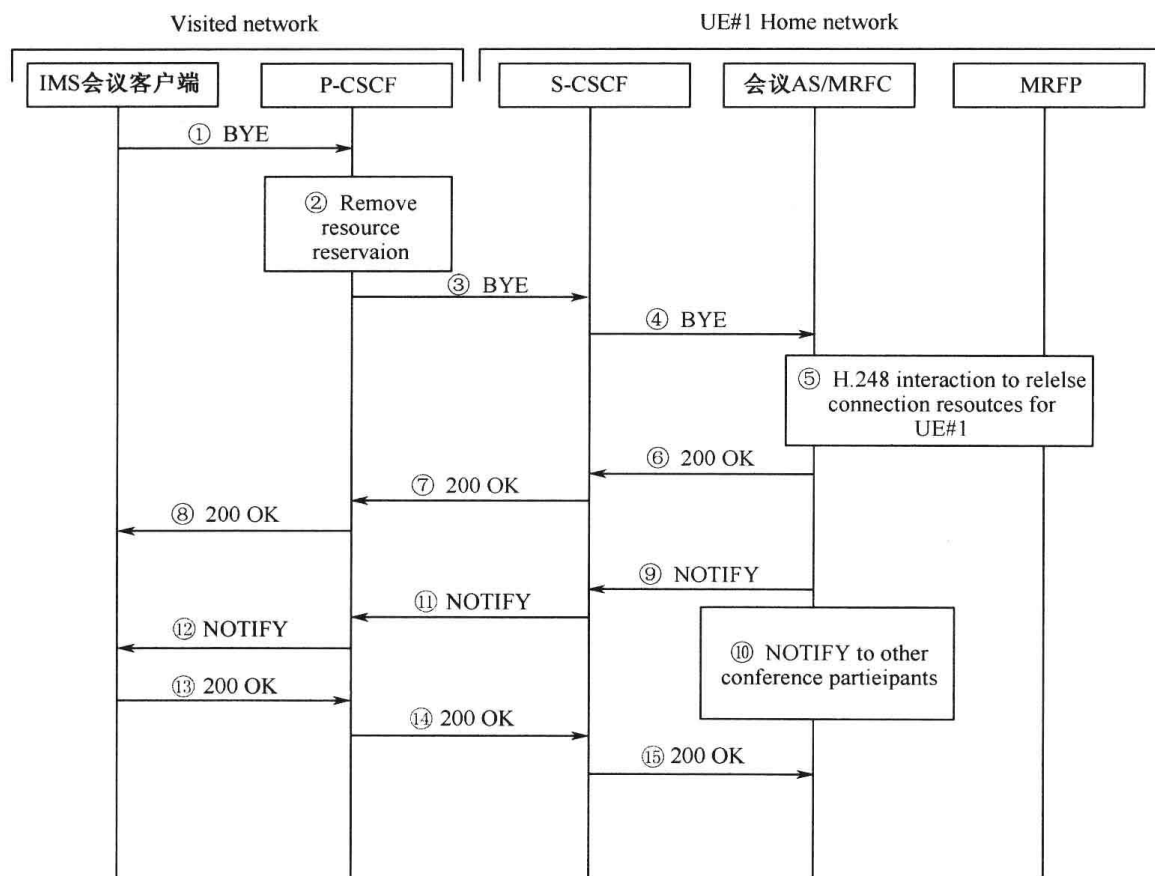


图 4.40 成员主动离席

5. 终止会议

当一个会议终止时，与之关联的会议策略和媒体策略也同时终止。任何对这些策略的读或写都将导致协议错误，并且会议 URI 变为不可用，任何发往它的 INVITE 消息或者 SUBSCRIBE 消息都将导致一个 SIP 错误响应。

在会议终止时，如果会议中还有成员，会议控制模块应该在真正终止会议之前向所有的成员发送 BYE 消息。同时，如果有用户订阅了会议通告服务，服务器也会在终止会议之前把订阅全部终止。

在 3GPP TS24.147 中，会议策略规定最后一个用户离开后会议即终止，那当最后一个离开后（使用 BYE 请求），会议就终止。在预约型的会议中，会议结束时间到也是一种会议结束的方式。

MC-5 接口的 HTTP 流程：在 3GPP 中提出了该接口上使用 HTTP 消息携带 CPCP（会议策略控制协议）消息体，进行用户与会议服务器的界面上交互。CPCP 使得会议的创建人可以进行如下控制：

- 定义会议的生存时间；
- 定义那些用户可参加会议，即访问控制列表（Access Control List, ACL）；

- 决定拨出和拨入列表；
- 定义会议中具有的角色和与这些角色相的责任；
- 发布通用的会议信息；
- 定义资源控制（Floor Control）策略。

可以采用 Web 应用的方式开发，具体信令不再赘述。

第5章 视频会议系统的测试方法

视频会议系统通过视频会议网络完成多点之间的音 / 视频信号的相互交互，最终实现身在异处的人们实现“面对面”的信息交流的需求。但视频会议又不同于一般普通的多点可视通信交流，它要求有一定的现场感和真实感。因此评价测试视频会议系统，应从人对视频会议系统实际感知体验效果出发。

影响参会者对视频会议实际感知的因素主要有：视频会议系统功能的完备性，是否可以模拟真实的会场情景；视频图像质量的好坏，人眼是否感觉舒适；视频会议系统设备硬件性能，包括节点设备性能和视频会议网络性能。

因此，完成对一个视频会议系统测试，应从下述三个因素具体展开。

① 使用者对视频会议系统实际感知体验，主要完成视频会议系统的功能测试。

② 视频图像质量评价，包括人眼主观评价和仪器客观测试。人眼是视频图像的最终接收者，只有主观测试满足预期的要求，用户才能真正获得满意的视频会议效果。测试仪器对视频图像质量的客观测试目前还不能替代主观测试，但可作为在视频会议系统选型中完成对厂商设备性能的横向比较的工具，同时也可兼作视频编码器的性能测试的工具。

③ 视频会议系统的硬件性能测试，主要完成对视频会议系统设备节点进行相关的电、光性能指标的评价，包括视频会议显示器的测试、视频会议网络 QoS 评价。

按照上述测试评价原则，本章重点描述下列内容：

- 视频会议系统功能测试研究；
- 视频会议图像质量测试研究；
- 视频会议显示器性能测试研究；
- 视频会议网络 QoS 性能测试研究。

5.1 视频会议系统功能测试研究

视频会议系统功能测试主要考察设备的业务能力、兼容性以及网络互通性，从实际使用的角度来完成对视频会议系统的总体评价。功能测试采用主观人为评价，为保障公平，将有多名测试者参加，并将各测试者的主观评分取平均值作为

最后评价结果。

测试环境在实验室内搭建，包括 GK、MCU、视频终端和以太网交换机，视频会议设备硬件及软件版本应与现网运营的版本相一致，测试拓扑如图 5.1 所示。

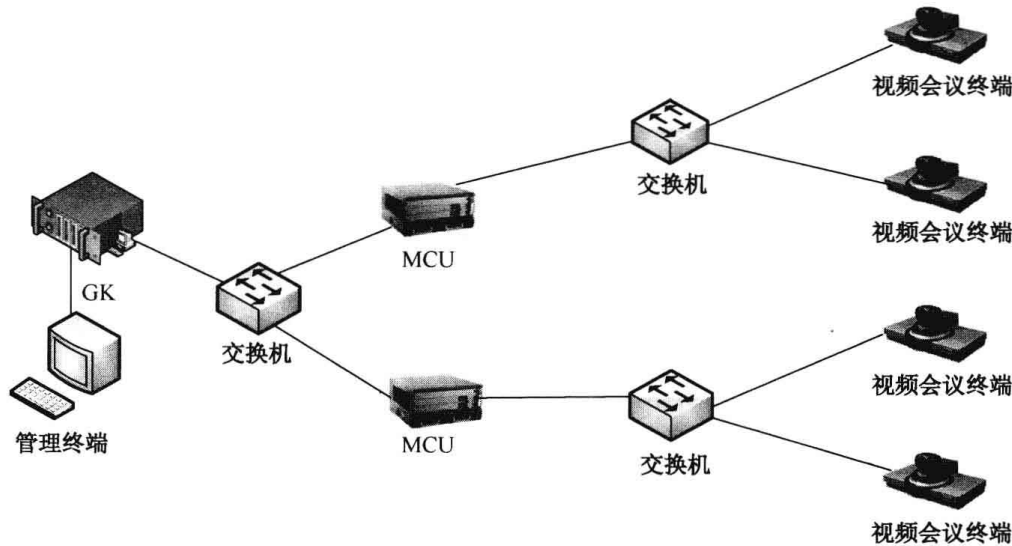


图 5.1 视频会议系统测试拓扑图

测试中主要选定业务常用的视频编码标准、图像分辨率、图像帧率、网络接入速率的组合格式，如表 5.1 所示。

表 5.1 常用的视频会议业务格式

视频编码标准	H.261		H.263		H.264					
图像分辨率	标清 CIF		标清 4CIF		高清 720P			高清 1 080P		
图像帧率	30 fps		30 fps		30 fps		60 fps	30 fps		
网络接入速率	384 kbps	768 kbps	768 kbps	1 Mbps	1 Mbps	1.5 Mbps	2 Mbps	1.5 Mbps	2 Mbps	4 Mbps

5.1.1 基本业务能力测试

基本业务能力测试主要针对同一厂商的网络设备和终端设备，其测试结果可作为视频会议系统设备选型的参考数据。

1. 音 / 视频效果测试

主要测试在 MCU 控制下（MCU 分别工作在编解码方式和转发方式下）各种会议类型下的音 / 视频效果，测试音 / 视频源分别取自摄像机现场输入或规定的节目源。

在测试过程中，多终端从表 5.1 中选取相同或不同的会议格式建立会议连接，

暂不发送双流信号，对会议的音 / 视频效果进行评估，并记录在测试记录表格中，如表 5.2 所示。

表 5.2 音 / 视频效果主观测试记录表

测试内容	测试者	终端 1（会议格式 1,2,⋯,n） 主观效果评分	终端 2（会议格式 1,2,⋯,n） 主观效果评分	终端 n（会议格式 1,2,⋯,n） 主观效果评分
视频色彩、清晰度、图像连续性	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
语音质量	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
音 / 视频同步	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			

2. 会场控制功能测试

主观评价会议控制过程中的广播、查看、轮询、分屏、静音、混音、语音激励等功能，记录在表 5.3 所示的测试记录表格中。

表 5.3 会场控制功能测试记录表

测试内容	测试者	会议格式 1 主观效果评分	会议格式 2 主观效果评分	会议格式 n 主观效果评分
广播功能 主会场广播分会场，所有分会场看到的是被广播会场的画面。观察各会场终端音视频质量，图像切换过程有无明显花屏	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
查看功能 主席会场可以在会议过程中浏览任意一个会场画面，而各分会场看到的是主席会场的画面。观察各会场音视频质量，图像切换过程有无明显花屏	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			

续表

测试内容	测试者	会议格式 1 主观效果评分	会议格式 2 主观效果评分	会议格式 n 主观效果评分
轮询功能 主席会场可以按照轮询的时间间隔依次浏览每个会场画面，而各分会场看到的是主席会场的画面。观察轮询时各会场音视频质量，图像切换过程有无明显花屏	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
分屏功能 主会场可以按照设置的分屏格式（如 2 分屏、4 分屏、8 分屏）同时查看多个会场画面，而各分会场看到的是主会场的画面，观察分屏后音 / 视频质量	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
静音功能 关闭全体终端麦克风或扬声器是否便捷；关闭某一终端麦克风或扬声器是否便捷	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
混音功能 两个或两个以上终端打开话筒，同时发言，在主会场听效果	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
语音激励 广播声音最大的会场画面，观察音视频效果	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			

3. 双流功能测试

视频会议系统的双流功能是指视频会议终端可以接收或发送多于一个的视频信道，通常在传输一路活动的人物图像主视频的同时，同时向远端会场传送 PC 信号，终端需要具备视频混合或切换的能力，以便向用户显示多路视频信号。

在测试过程中，MCU 分别工作在编解码方式和转发方式下，终端从表 5.1 中选取相同或不同会议格式建立会议连接，并同时发送主音 / 视频信号和 PC 流信号，对会议的音 / 视频效果进行评估，并记录在测试记录表格中，如表 5.4 所示。

表 5.4 音 / 视频效果主观测试记录表

测试内容	测试者	终端 1 (会议格式 1,2,...,n) 主观效果评分	终端 2 (会议格式 1,2,...,n) 主观效果评分	终端 n (会议格式 1,2,...,n) 主观效果评分
主视频色彩、清晰 度、图像连续性	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
语音质量	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
PC 图像信号质量	测试者 1			
	测试者 2			
	测试者			
	单项平均分			

4. 网络环境适应能力测试

网络环境适应能力测试模拟视频会议系统在现网环境下对网络丢包、延时、抖动的适应能力，以及对抗丢包、抗抖动的处理效果。网络丢包率、延时和抖动的效果可通过串接在测试环境中的网络仿真仪来实现，具体测试过程如下所述。

- 将网络仿真仪串接在 2 台待测终端之间（终端之间通信采取不通过 MCU 和通过 MCU 两种），根据现网的 QoS 质量，在仿真仪上分别设置网络丢包率、延时和抖动。从表 5.1 中选取某一会议格式，召开一个视频会议。
- 记录音 / 视频效果（色彩、清晰度、图像连续性、图像停顿时间间隔；声音连贯性、清晰程度、声音停顿时间间隔等），注意观察各厂家对抗丢包、抗抖动的处理效果（如降帧率等）。
- 设置不同的网络丢包率、延时和抖动，继续记录在此条件下的音 / 视效果，直至语音和画画完全不可接受。

5.1.2 兼容与互通性测试

本测试项主要针对在不同厂商的网络设备和终端设备、不同的音 / 视频编码标准、不同的高标清图像分辨率、不同的会议终端网络接入速率相混合的会议环境下，视频会议系统的实际兼容性和互通性，是评价已存视频会议系统的重要主观参考。

在测试过程中，MCU 分别工作在编解码方式和转发方式下，对于多厂商终端从表 5.1 中选取不同会议格式建立会议连接，并同时相互发送主音 / 视频信号和 PC 流信号，测试者对会议的音视频效果进行评估，记录在测试记录表格中，如表 5.5 所示。

表 5.5 兼容与互通性主观测试记录表

测试内容	测试者	A 厂商终端 1 (会议格式 1,2,...,n) 主观效果评分	B 厂商终端 2 (会议格式 1,2,...,n) 主观效果评分	C 厂商终端 n (会议格式 1,2,...,n) 主观效果评分
观察待测终端能否顺利入会，会议过程中有何种异常现象	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
视频色彩、清晰度、图像连续性	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
语音质量	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			
PC 双流信号质量	测试者 1			
	测试者 2			
	测试者 n			
	单项平均分			

5.2 视频会议图像质量测试研究

在视频会议系统中，视频处理过程的各个环节将对视频图像质量产生不可避免的负面影响，如编码器对视频内容的压缩、解码器对视频内容的解压缩、格式转换对视频分辨率和高宽比的影响、IP 网络对视频流造成的延迟和产生的额外噪声等。

在现有的 H.263、H.264、MPEG 系列、AVS 等图像压缩编码标准中，编码器的计算复杂度一般是解码器的 5~10 倍，任何一种压缩编码器势必对图像质量造成压缩损伤。实际上，在整个视频传送过程中，只有编码器对图像质量的影响最大，因此对图像质量的测试也间接完成了对图像编码器的评价。对于解码器的测试，主要是考察其输出的电气性能指标，由于解码器相对编码器并不是影响

视频图像质量的关键因素，因此对解码器的测试不作为视频会议系统测试的组成部分。

目前，图像质量的测试可分为主观测试和客观测试。由于人眼常常是图像的最终接收者，因此人眼的判断也许是图像质量评价最有效的手段。事实上，一开始的图像质量评价确实依赖于这种主观评价手段，而且主观评价方法和标准也相对较完善。但其缺点是无法应用数学模型对其进行描述，不易定量准确测量，易受客观因素的影响，而且比较费时费力；另外，主观质量评价的随意性很大，同一个人在不同情况下对同一幅图像可能会评出不同的分值。随着图像技术的发展，人们希望能够用客观的和可定量分析的数学模型衡量出人们对图像的这种主观感受，即客观图像质量评价方法。客观图像质量评价的目标就是设计能够自动预测人眼对图像和视频的感知质量的方法。

人们希望对客观图像质量的评价结果与人的视觉感知相一致，也就是说，客观评价的结果应该尽量符合人们对图像质量的主观真实感受，因而在各种评价方法中或多或少地引入了人眼视觉特性。然而在机器视觉领域，至少目前还无法设计出与人的复杂视觉感官相一致的系统，更多的是设计出定量模型及参数以供计算机处理。

因此，人为的主观评价方法是视频会议图像质量测试的主要方法，仪器仪表的客观测试可作为不同厂商设备图像质量的横向比较方法，但是只作为参考，不作为最终结果评定。

5.2.1 视频图像质量主观测试

主观测试就是通过直接收集用户对视频质量的感受，给出视频质量评价的方法。评价用户体验（QoE）最可靠的方法是主观评价，因为在大多数情况下，人是视频最终的接收者，人的主观感觉直接反映了人对视频图像质量的感受，所以人的主观感受是最可靠的评价标准，由一系列观测员打分获得的主观视频质量评价方法一直是最可靠的质量评价方法。视频质量的主观评价法就是选择一批非专业的观测员在一个受控的环境中连续观看一系列测试视频序列，然后采用适当的评价方法让他们按照自己的主观感受对视频序列的质量进行评分。

近年来 ITU（International Telecommunication Union，国际电信联盟）先后推出了多个针对多媒体应用的主观质量评估标准，如 ITU-R BT.500-11、ITU-T P.910、ITU-T BT.1788 等，提出了标准的可视条件、观察者的选择、测试内容的标准、评估程序以及数据分析的方法。

目前比较常用视频质量的主观评价法有两种，一种是单刺激主观评价法；另一种是双刺激主观评价法。单刺激主观评价方法不需要原始视频序列，主要包括

ACR (Absolutely Category Rate) 方法; 双刺激主观评价法可以有选择地采用原始视频序列与失真视频序列对比或者在不同类型的失真视频序列间进行对比评分的方法, 主要包括 DSCQS (Double-stimulus Continuous Quality-Scale)、DSIS (Double-stimulus impairment scale) 等方法。

这两种方法也可以简单看成等级评价方法和比较评价方法, 无论哪种方法, 都包含一种 5 级制的尺度评价标准, 在单刺激主观评价法中, 主要是直接对处理过的视频进行视频质量等级的优、良、中、差、劣 5 级主观评分; 在双刺激主观评价法中, 主要是对测试视频相对于原始视频质量损伤感受来进行 5 级评分, 基本评分标准如表 5.6 所示。

表 5.6 视频图像质量主观评价

评分	视频质量等级	视频损伤感知等级
5	优 (Excellent)	没有察觉
4	良 (Good)	有察觉但不讨厌
3	中 (Fair)	察觉且稍微讨厌
2	差 (Poor)	讨厌但不令人反感
1	劣 (Bad)	极讨厌且令人反感

此外, 还有一种基于一个或多个视频质量要素的属性指标评价方法, 其指标主要有流畅性、亮度、对比度、清晰度、抖动程度、音质等。观测员对每个测试视频序列的上述各项属性指标打分, 再得到一个总分数, 对所有观测员的评分求平均值, 得到最终的质量评价结果。

1. 单刺激主观评价法

单刺激主观评价法中主要采用 ACR (Absolutely Category Rate, 绝对尺度分级法), 它以与视频质量有关的 5 个形容词“优、良、中、差、劣”作为基准, 来衡量评价观测人员在的主观反映, 一般分为 5 段、9 段和 11 段进行打分, 分别称为 ACR5、ACR9 和 ACR11, 其中 ACR11 增加了两个等级, 即“特别优秀”和“特别差”。

在 ACR 测试方法中, 每次给观测员观看一个测试视频序列, 然后让观测员独立进行评分。图 5.2 为 ACR 方法的示意图, 其中图像 A_i 表示经过 i 类型的处理后的视频图像序列 A, B_j 表示经过 j 类型的处理后的视频图像序列 B, C_k 表示经过 k 类型的处理后的视频图像序列 C, 两个视频序列播放之间的时间用于测试人员评分。

2. 双刺激主观评价法

双刺激主观评价法是一种比较评价法, 目前较为常用的是 DSCQS (Double-

stimulus Continuous Quality-Scale, 双激励连续质量分级法), 该主观测试需要一个特殊的群体, 而且需要经过必要的训练之后, 在一定要求的环境下, 对被测视频序列和标准视频参考序列进行比较, 以发现被测序列的损伤。

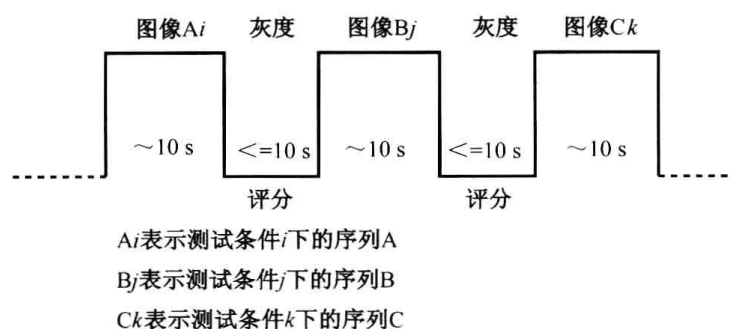


图 5.2 ACR 方法示意图

首先, 在评价之前对观察者需要有一定的培训指导, 向他们显示什么样的图像效果是 ‘Good’ 或 ‘Poor’, 等等。两个不同版本 (A 序列和 B 序列) 而相同内容的视频片段被重复播放两次 (总计 4 次播放), A 和 B 视频序列可能是参考序列, 也可能是经过图像处理后的被测序列。呈现给观察者眼前的视频版本和序列性质完全是随机的, 观察者之前并不知晓。观察者需要对 A 和 B 两种序列给出连续的评级, 过程如图 5.3 所示。

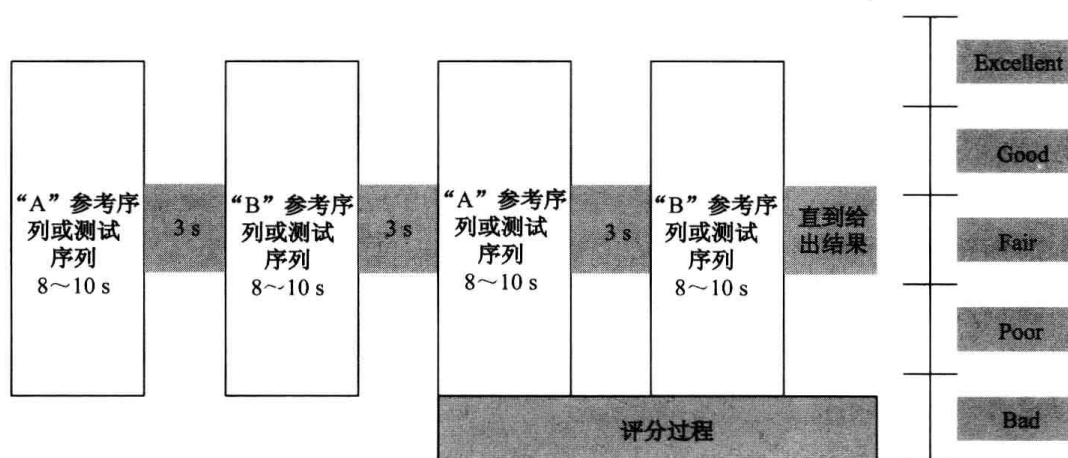


图 5.3 DSCQS 评分过程

图像质量的评分以每个观察者的 MOS (Mean Opinion Score) 主观意见评分的平均数为基础来进行, 将被测序列的 MOS 评分减去参考序列的 MOS 评分, 得到 DMOS 评分 (Differential Mean Opinion Score)。如果 DMOS 分值为正, 说明被测序列视频质量好于参考序列; 如果为零, 说明两者质量一致; 如果为负, 说明被测序列视频质量低于参考序列。

3. 属性指标评价方法

ITU-T P.910 特别定义了一种称为指标测试视频主观质量测试方法，它是一种更关注细节的评价方法，当利用单刺激或双刺激主观评价方法所得的分数很接近的时候就可以采用对视频图像进行属性指标的测试来进行更为细致的评价，从而得到更准确地评价结果。ITU-T P.910 定义的属性指标如表 5.7 所示。

表 5.7 属性指标

Brightness	亮度
Contrast	对比度
Color reproduction	彩色复制
Outline definition	轮廓清晰度
Background stability	背景稳定性
Speed in image reassembling	图像重组速度
Jerkiness	抖动
“Smearing” effects	拖尾效应
“Mosquito” effects	蚊子效应
Double image/shadows	双图像/阴影
Halo	光环

以上的指标大部分是由于数字压缩编码所带来的图像质量损伤，并且很多指标用主观评价方法很难对其评价。将上面提到的参数进行分类，进而归纳为几个典型参数，如流畅性（反映观看视频时的流畅程度）、清晰度、画质（反映视频图像的亮度、色彩、对比度和锐化程度等）、抖动程度和同步性这 5 项基本属性指标，观测员在观测视频序列时，通过评价这些指标的分数来评价视频的总体质量。

5.2.2 视频图像质量客观测试

主观视频评价方法的优点是测试结果比较直接，而且已经被长期应用在各种视频系统图像质量的评价中。其缺点是耗时长、对评价人员有一定的要求，评价结果与评价人员的个人喜好、视频显示内容、测试人员身体疲劳程度和所经过的测试培训过程有很大的关系。因此，视频图像质量客观测试方法也日益受到关注，虽然目前客观测试结果尚不能完全真实反映实际结果，但一直是人们努力的方向。另外，客观图像视频质量评价使用仪器仪表来进行，适合重复性比较强的测试，图像质量客观测试结果可作为不同厂商的设备性能横向对比的参考数据。

图像质量的客观评价方法是使用一个或多个图像的度量指标来反映图像的质

量情况, 目前各种算法大致有以下几种分类。

- 从参考源可用性角度可分为: 全参考质量评价、部分参考质量评价和无参考质量评价。
- 从评价处理方式可分为: 空间域图像质量评价方法、频域图像质量评价方法、空间域和频域综合的图像质量评价方法。
- 从评价指标角度可分为: 单因素质量评价(噪声、模糊、块效应等)和综合因素质量评价方法。
- 从视觉心理生理角度可分为: 自顶向下的方法结合、结合人眼视觉系统的自底向上的方法。
- 从应用智能角度可分为: 基于神经网络、机器学习、模糊理论等。

在上述这些分类方法中会有部分交叉, 并没有清晰的界限。目前比较常见和应用较多的是从参考源可用性角度定义的测试方法, 分为全参考测试、无参考测试和部分参考测试。全参考测试是对参考序列和被测序列进行像素对像素的完全比较, 无参考测试仅对被测序列进行测量, 部分参考是对参考序列和被测序列的部分提取指标进行测试。

在这三种客观质量评价方法中, 全参考评价方法是相对比较成熟的, 也比较与主观测试评价相一致, 但它必须借助原始的非失真图像作为参考, 这极大地限制了它的应用范围, 如在无线传输等评价场合较难满足。

全参考的评价方法分为两类: 第一类是在图像噪声层面上, 基于全像素失真统计的传统评价方法; 第二类是基于人类视觉系统的评价方法。

1. 基于全像素失真统计的传统评价方法

此类方法以统计理论为基础, 逐帧、逐像素比较参考源和测试源的数据差异, 通过获得原数据和待测数据间的总体误码累计, 体现像素噪声层面上的绝对误码率, 从而反映图像的视频质量。该类评价方法能敏感捕获两幅图像在像素层面上的细微失真, 具有很高的敏感性。此外, 其数学方法简单, 物理意义清晰, 是目前应用最广的评价方法, 其模型如图 5.4 所示。

该类传统的全参考评价方法主要代表是 PSNR (Peak Signal-To-Noise Ratio) 测试, 主要是比较参考序列和被测序列的信号差值的信噪比。PSNR 的数学表达式为

$$\text{PSNR} = 10 \lg \left[\frac{(2^n - 1)^2}{\text{MSE}} \right] \quad (5.1)$$

式中, MSE 表示原始图像和解码图像像素间的均方差, n 表示图像像素值所用的比特数, 通常, $n=8$ 。PSNR 的单位为分贝 (dB)。解码图像的 PSNR 值较高

(MSE 值较低), 表示其客观质量较高; 相反, 解码图像的 PSNR 值较低 (MSE 值较高), 表示其客观质量较低。

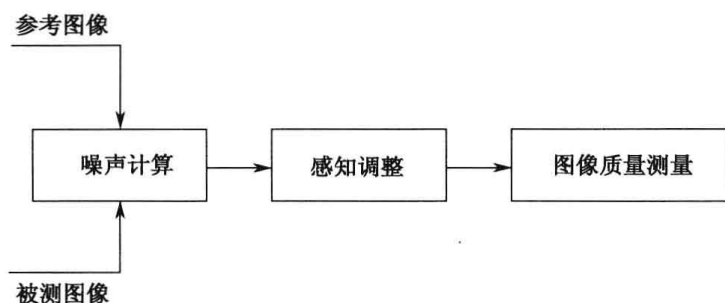


图 5.4 基于全像素失真统计图像质量评价模型

该方法虽然从总体上反映原始图像和恢复图像像素层上的差别, 但由于人类的视觉系统对高频空间特性的敏感度比低频空间特性低, 因此仅对被测图像和参考图像的图像之间的信号差值进行测量所得到的结果并不能真实反映图像的人眼感知质量, 其评价机制无法体现出图像视频数据所承载的不同于一般数据的内容信息, 在有些情况下, 甚至是完全相反的结果, 我们通过图 5.5 为例进行说明。

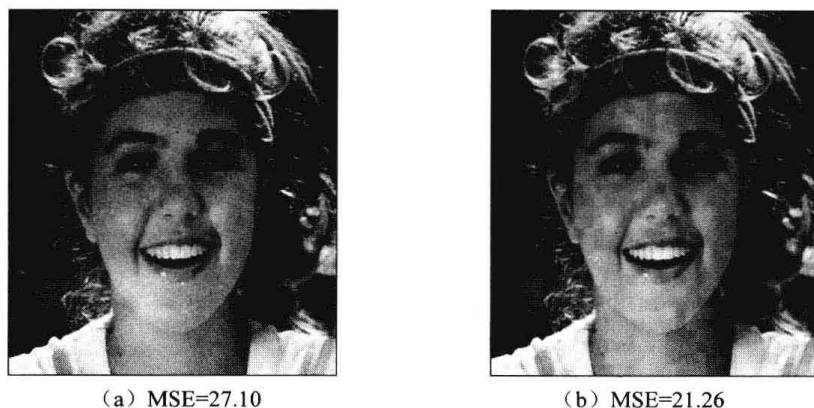


图 5.5 PSNR 测试结果分析

在图 5.5 中, 我们通过 PSNR 方法分别计算原始画面与处理后画面的均方差, 可以看到画面 5.5 (a) 的均方差大于画面 5.2 (b) 的均方差, 如果据此判定, 则画面 5.5 (b) 的图像质量应好于画面 5.5 (a)。

但实际人眼判定的效果是画面 5.2 (a) 的质量好于画面 5.2 (b)。产生这种结果的原因是画面 5.2 (a) 中包含了较多的高频分量错误, 而画面 5.2 (b) 包含了较多的低频分量错误。由于人眼对低频分量较高频分量比较敏感, 因此即使基于噪声的 PSNR 判定画面 5.2 (a) 的质量较差, 但人眼仍认为画面 5.2 (b) 的质量较好。

可见, 人眼的视觉感知结果并不与画面的信噪比大小相一致, 人眼自身的“生理特点”和人关注内容的“心理特点”都对绝对误码效果产生不同程度的掩蔽

效果,使得对图像质量优劣的理解并不仅仅依赖绝对的误码率,因此在基于全像素失真统计的传统评价方法之上,人们又提出了基于 HVS (人类视觉系统) 的评价方法。

2. 基于人类视觉系统的评价方法

当前,基于 HVS 的评价方法主要可以划分为基于视觉感知的算法模型和基于视觉兴趣加权的算法模型两类主流算法模型。

基于视觉感知的算法模型在整个评价过程中综合考虑了人眼的视觉非线性、对比度敏感性、多通道特性、掩盖效应等因素,通过误差综合模拟人脑对来自各个视觉通道的神经信号的总体处理和认知结果,其最终的结果与主观评价结果较吻合,其模型如图 5.6 所示。

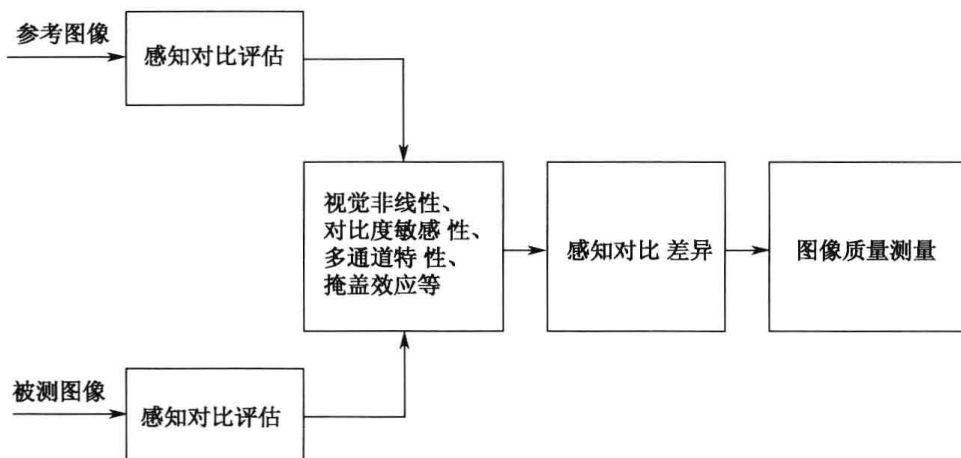


图 5.6 基于 HVS 图像质量评价模型

此类基于人类感知系统的方法如 PQR 测试 (Picture Quality Rating), 在仪表测试过程中, 其测试参数的配置充分参考了主观人眼测试对观察环境和显示设备的要求, 因此该测试结果与人类的主观感知效果比较接近。PQR 测试基于 ITU-T J.144 标准《全参考图像下数字有线电视客观感知视频质量的测量技术》, 重点考察被测图像与参考图像之间的差异。

测试过程采用两个视频文件作为测试依据: 一个文件是原始的参考视频序列, 另一个文件是该参考视频序列经过压缩、损伤或经处理过的序列版本。首先, 图像质量分析仪对这两个视频序列在空间域和时间域中进行校准, 而后仪器对基于人类视觉系统和提示模型的被测视频质量进行分析, 输出与主观评测相关的质量测量结果。

基于视觉兴趣加权的算法模型适用于提取兴趣区域 (Region Of Interest, ROI) 和场景内容较单一的场合。这类算法模型的核心是确定兴趣区域的划分以及权值大小的选择。目前 ROI 判别思路主要有频域变换和边缘滤波两种思路。频域

变换后所得的子带与 HVS 兴趣特性非常近似,可根据此特点对较重要的域失真给予大的加权系数体现;边缘滤波则可有效提取人眼较为敏感的区域信息,通过合理选择滤波的阈值门限,也可达到兴趣区的有效划分。权值大小选择则需要根据实测修正,不断总结并选择合适的经验权值。从总体上看,基于 ROI 加权的算法模型在如何合理划分兴趣域及如何调整权值等关键技术方面还有待进一步研究。

总之,基于 HVS 的图像质量评价模型是在模拟人眼视觉系统的基础上提出的,与主观测试结果一致性较好。值得注意的是,在严格的测试条件和不同的图像失真情况下,没有任何一种现有的视觉感知方法比简单的数学方法如 PSNR 表现出明显的优势。实验也表明,在某些场合如非线性的情况下,PSNR 方法的结果甚至超出了 HVS 方法。但是由于目前对于人眼视觉系统的研究还不是很透彻,人眼视觉系统的数学模型并不能够真实反映人眼的主观感受,开发通用的基于 HVS 的特性的评价方法仍是目前研究的一个难点。

5.3 视频会议显示器性能测试研究

视频会议显示器是视频会议系统中最直接影响观看者满意度的设备,随着全球数字电视技术的不断发展,显示技术也不断推陈出新,例如,采用宽色域技术提高图像的鲜艳度,提高帧帧的“插灰”技术降低图像的运动拖尾,LED 背光技术使显示器变得节能、超薄等,这些技术大大地满足人们对于色彩、高清晰度图像等方面的要求,同时相对于传统电视显示器也更加轻薄,辐射更小,更加节能和安全。

国际上在平板显示领域有 IEC、ISO、VESA 等组织制定的技术标准,我国参考国际有关的标准并根据国情制定了数字电视显示特性的测量标准,其中涉及数字电视平板显示器性能标准的有 5 项: SJ/T 11339-2006《数字电视等离子体显示器通用规范》、SJ/T 11343-2006《数字电视液晶显示器通用规范》、SJ/T 11348-2006《数字电视平板显示器测量方法》、SJ/T 11325-2006《数字电视接收及显示设备可靠性试验方法》、SJ/T 11326-2006《数字电视接收及显示设备环境实验试验方法》。

数字电视显示器相关技术标准和测试标准,完全可以适用于视频会议系统显示器性能的测试研究。

视频会议显示器的测试项目包括亮度、对比度、亮度均匀性等,本节将针对其主要性能和测试方法进行分析和介绍。关于显示器更加详细的性能测试方法,有兴趣的读者可参考有关专门文献。

对显示器的测试主要通过标准视频信号发生器,产生模拟和数字的标高清标准测试信号,将其输入进显示器,通过亮度计和色度计来测试显示器的亮度和色

度指标，测试连接框图如图 5.7 所示。

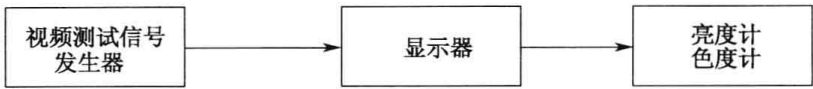


图 5.7 显示器测试系统框图

测量亮度、色度时应在暗室中进行，暗室内应没有可引起光线反射的物体，要求暗室杂散光照度 $\leq 1\text{ lx}$ ，以避免对测试结果的影响。

测试前需要进行准备工作，显示器的工作状态应调整到出厂设置状态或最佳状态，光学测试仪器的光轴应与显示器中心区域正交垂直，其距离为屏幕高度的 3~4 倍。

5.3.1 亮度测试

1. 亮度的概念

根据 2.1.2 节的描述，光亮度是表示发光面明亮程度的，指发光表面在指定方向的发光强度与垂直且指定方向的发光面的面积之比，单位是坎德拉 / 平方米 (cd/m^2) 或尼特 (nt)。

对于显示器或电视机的亮度，根据采用不同的调整状态和测试信号，有以下 4 种表示屏幕亮度的性能指标。

- ① 有用峰值亮度：以白窗口信号作为测试信号，显示器预设正常对比度和亮度，用亮度计测试白窗口的亮度值称为有用峰值亮度。
- ② 有用平均亮度：以 100% 的全白场信号作为测试信号，显示器预设正常对比度和亮度，用亮度计测试的屏幕中心位置的亮度值称为有用平均亮度。
- ③ 最大峰值亮度：以白窗口信号作为测试信号，显示器预设最大对比度，用亮度计测试的白窗口的亮度值称为最大峰值亮度。
- ④ 最大平均亮度：以 100% 的全白场信号作为测试信号，显示器预设最大对比度，用亮度计测试的屏幕中心位置的亮度值称为最大平均亮度。

上述几种亮度值代表了平板显示器亮度值的不同概念，对于正常观看电视图像而言，有用平均亮度更有意义，因此，在我国行业标准中，只规定了有用平均亮度。

显示器的亮度设计要考虑人眼的实际需求，人眼视力与环境条件密切相关，当亮度为 $0.1\text{ cd}/\text{m}^2$ 时，视力约为 0.6。当亮度增加到 $1\text{ 000 cd}/\text{m}^2$ 时，视力达到 2.3 左右，之后亮度再增加视力不会明显增大了，而且当亮度过大时，就会感到耀眼，使图像清晰度下降，甚至无法睁眼。大量统计表明，在正常人的视力情况下，平板显示屏的亮度、对比度、色饱和度适中，图像柔和、清晰，层次分明，

这时的有用平均亮度为 $60 \sim 70 \text{ cd/m}^2$ 。

2. 亮度的测试方法

按照图 5.7 所示连接测试系统, 测试信号发生器分别输出全白场信号和白窗口信号。按照下列步骤进行测量。

- ① 显示器预设正常对比度和亮度。
- ② 输入白窗口信号作为测试信号, 用亮度计测试如图 5.8 中心 P_0 点亮度值, 得到有用峰值亮度。
- ③ 输入 100% 全白场信号作为测试信号, 用亮度计测试如图 5.8 中心 P_0 点亮度值, 得到有用平均亮度。

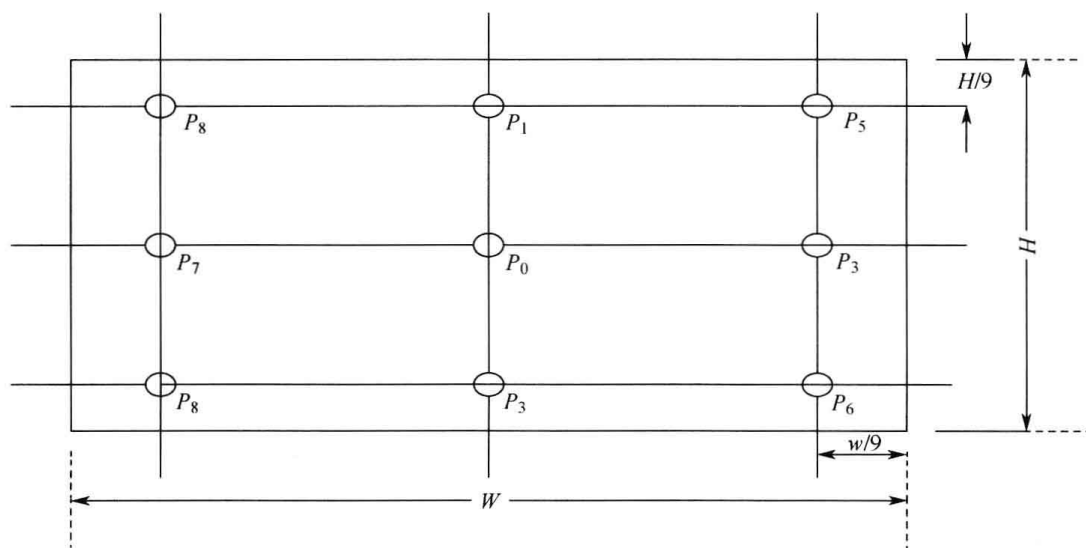


图 5.8 显示器亮度测试点

3. 亮度均匀性概念和测试

亮度均匀性是屏幕中心亮度与屏幕边缘之间的亮度差, 它表征了屏幕中心与屏幕其他位置, 尤其是屏幕边缘所显示亮度均匀性的程度。

亮度均匀性按照以下公式计算:

$$P_i = \frac{L_i}{L_0} \times 100\% \quad (5.2)$$

式中, L_0 是屏幕中心点的亮度值, L_i 是偏离中心点的亮度值。如果亮度均匀性特别好, 则 L_0 和 L_i 的亮度值相等, 则 $P_i=1$, 可见, 亮度均匀性越大, 表明该显示器的亮度均匀性越好, 均匀显示各点亮度的能力强。

测试亮度均匀性时, 首先将显示器调整到规定的标准工作状态。将全白场信号输入到显示器, 用亮度计测量如图 5.8 中 $P_0 \sim P_8$ 各个点的亮度值, 分别记为

$L_0 \sim L_8$ ，然后，按式（5.2）计算得到亮度的均匀性。

不同显像原理的显示设备亮度均匀性各有差异，对于 CRT 显示器，受电子束偏转的影响，其屏幕中心亮度最高，屏幕边缘亮度较其中心亮度会降低，因此 CRT 显示器的亮度均匀性会比较低。对于 LCD、PDP 显示器，由于像素点与 RGB 激励信号是点对点对应显示，因此其亮度均匀性要好于 CRT 显示器。

5.3.2 对比度测试

1. 对比度的概念

对比度是表征显示器重显图像灰度的能力，对比度高代表重显的图像层次多。适当的图像对比度可以使图像层次分明，观察图像时有纵深感。关于对比度概念的一般说明可参考 2.2.5 节。

2. 对比度的测试方法

将显示器的对比度和亮度调整到正常状态下，把如图 5.9 所示的黑白窗口测试信号加到显示器，测试黑窗口的亮度值 L_1 、 L_2 、 L_3 、 L_4 和白窗口的亮度值 L_0 ，则对比度使用式（5.3）进行计算。

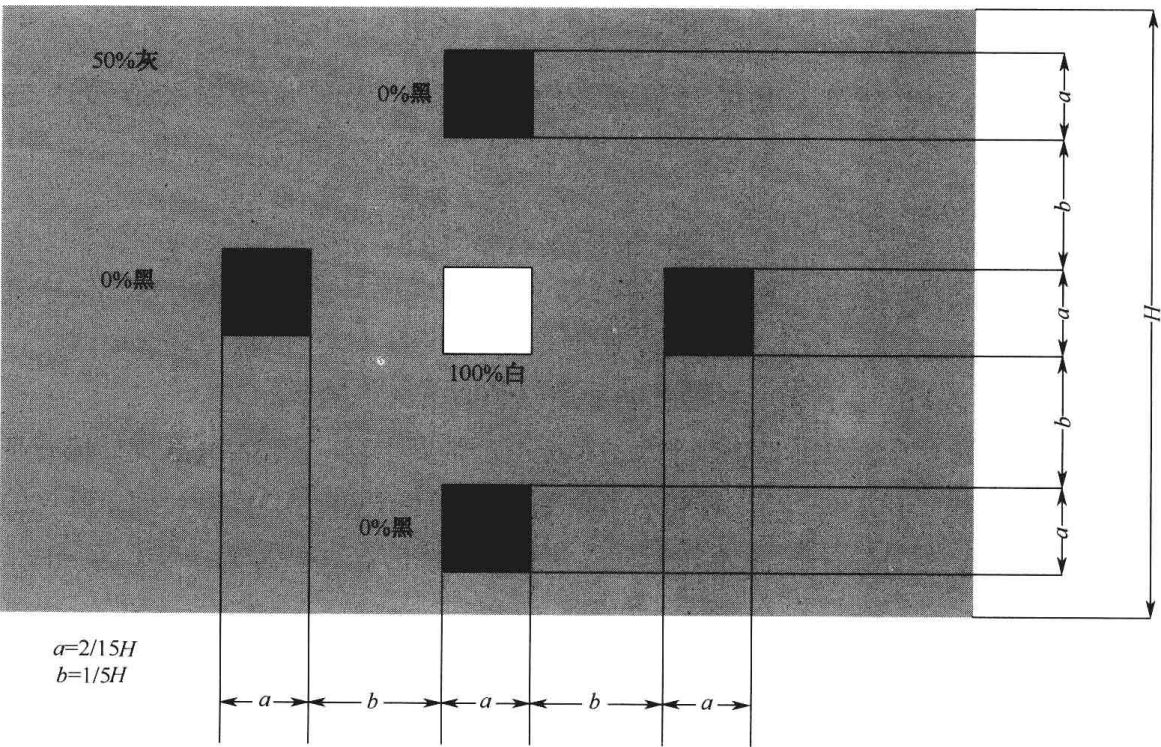


图 5.9 黑白窗口测试信号

$$C = \frac{4L_0}{L_1 + L_2 + L_3 + L_4} \quad (5.3)$$

在测试黑窗口的亮度时，要使亮度计对准黑窗口的中心点。

5.3.3 色域覆盖率测试

1. 色域覆盖率的概念

色域覆盖率表示在 1976 CIE 均匀色度空间坐标系 (u' , v') 中，对基色（红、绿、蓝）所对应的显色三角形的面积度量，它代表了显示器的彩色还原能力。

根据 2.3 节的介绍，利用色度图可以确定某种颜色在色度图上的色度坐标，同时，在色度图上标出 RGB 三基色坐标点的位置，画出由它们构成的三角形，三角形的面积代表了由这些基色所能配出的彩色范围。色域覆盖率即为显示器的色域面积（即三基色 RGB 三角形的面积）占 u' 、 v' 色度空间全部面积的百分比，该数值越大，说明显示器重现还原的彩色越多，颜色越鲜艳。

按 1976 CIE-UCS 均匀色度空间的色度坐标 (u' , v')，用以下公式计算色域三角形的面积及色域覆盖率 C_p ：

$$C_p = \frac{S_{\text{RGB}}}{0.1952} \times 100\% \quad (5.4)$$

式中，色域面积 $S_{\text{RGB}} = \frac{1}{2}[(u'_R - u'_B)(v'_G - v'_B) - (u'_G - u'_B)(v'_R - v'_B)]$

常数 0.1952 是 1976 CIE- u' 、 v' 色度图中全部色域（光谱从 380~780 nm）的面积。

2. 色域覆盖率的测试方法

将显示器调整到规定的标准工作状态，把全红场信号、全绿场信号、全蓝场信号分别加到显示器设备上，用色度计依次测量如图 5.8 的 P_0 坐标 (u'_R, u'_B)、(u'_G, u'_G)、(u'_B, u'_B) 用式 (5.4) 计算色域三角形的面积及色域覆盖率。

按照上述测量方法测得 NTSC 制 C_白 和标准清晰度电视 (SDTV) 系统显像三基色的色域覆盖率为 38.1%；PAL 制和高清晰度电视 (HDTV) 系统 D₆₅ 显像三基色的色域覆盖率为 33.5%；我国行业标准规定：平板显示器的色域覆盖率应大于 32%。

5.3.4 清晰度测试

1. 清晰度的概念

图像清晰度是显示器上人眼观察图像清晰、细腻程度的标志，是显示器指标

中最为主要的综合指标之一，也是消费者追求的主要图像质量指标。

图像清晰度与图像分辨率有关，但图像分辨率并不等于图像清晰度。清晰度是指人眼看到的图像的清晰程度，是由系统和设备的客观性能的综合因素造成的人们对最终图像的主观感觉。虽然是主观感觉，但这种主观感觉是可以定量测量的，即可用黑白相间的线条间隔来衡量，并有标准的测试方法和测试图，其测量数据有明确的单位，即电视线。而分辨率与清晰度不同，它不是指人的主观感觉，而是指在摄像、录制、传输过程中所使用的图像质量指标，以及显示设备自身具有的表现图像细致程度的固有显示格式，即单幅图像信号的采样格式和显示设备的显示格式。图像信号的采样格式和显示设备的显示格式，均是用“水平像素×垂直像素”来表示的，其单位不是电视线数，而是像素点数。图像信号的分辨率和显示设备的分辨率是由制式和产品规格决定的，是固定不变的，而清晰度是因重现条件而可改变的。在诸如 LCD、PDP 等固定分辨率的平板显示器中，图像信号的源分辨率等于屏幕的几何显示格式，实现信号源与显示屏之间点对点显示，信号通道的带宽又足够宽，这时图像的清晰度可能达到最高值。例如，在 LCD、PDP 等 HDTV 电视显示器中，信号源的图像分辨率为 $1\,920 \times 1\,080$ ，经理想的图像格式变换器，在 $1\,920 \times 1\,080$ 物理格式的显示屏上，水平、垂直清晰度有可能达到 1080 电视线，关于图像清晰度的定量说明，可参考 3.2.2 节。

2. 清晰度的测试

将显示器调整到标准工作状态，将复合测试图信号输入显示器，观察显示器水平和垂直方向可分辨的明暗交替楔形线条的总数，即电视线，使用垂直楔形线观测水平清晰度，水平楔形线观测垂直清晰度。

5.4 视频会议网络 QoS 性能测试研究

视频传送对 IP 传送网络的要求就是要有 QoS 保障，只要满足延迟小、抖动低、丢包率低三个基本要求，视频传送的质量能得到较好的保证。

网络延时主要取决于网络拓扑和编解码设备的性能，如果延时过大，用户的主观体验将明显变差；网络抖动取决于网络的动态路由，以及网络设备由于拥塞造成的延时等因素，一般的视频解码器主要针对稳定的码流，如果经过线路传送后码流抖动较大，解码器对视频包到达的均匀性要求得不到满足，当超过解码器的承受范围时，解码器丢弃视频包，最终表现为丢包；网络丢包率取决于网络流量，当网络流量达到最大值时，由于网络设备的拥塞导致网络丢包。丢包率大，表明网络带宽不够，会直接影响到视频质量。

实际上,影响视频传送质量的主要因素可以归纳到网络丢包率的大小,至于影响的程度就要视编解码器的处理而不同了,容错性好的编解码器解出来的图像用户可能感觉不到有丢包,编解码器容错性差或编码器没有容错性,解出来的图像效果就会感觉到有缺陷,如马赛克、图像拖尾、停顿等。影响的大小同时还与视频压缩算法有关,有的算法压缩比大,丢一个包,信息丢失量也很大,对视频影响也大。此外,如果丢失的包是I帧,丢图像质量的影响就会更大,反之则小。根据实测结果,1%的丢包率对视频的影响不是很大;2%的丢包率会导致视频效果就比较差,但还可以勉强接受;3%以上的丢包率会导致明显的马赛克、图像拖尾,图像无法接受。

IP网络作为一个不稳定的网络,保证视频可靠传送的技术最理想的方法是保证网络的可用带宽,网络可用带宽取决于网络的设计带宽,网络的背景流量等因素,当网络繁忙期间,网络可用带宽自然降低。因此在规定的传送带宽要求下增加一定的富裕量,是降低丢包率的有效方法。

对于实际IP视频网络的性能测试,现有的软件工具比较多,如IPSwitch公司的WSPing等,可以有效地完成网络时延、抖动、丢包、网络可用带宽指标测试和丢包发生位置探测。

5.4.1 网络时延的基本概念和时延测量原理

时延是指一个报文或分组从一个网络的一段传送到另一端所需的时间。时延由发送时延、传播时延和处理时延组成。发送时延是节点在发送数据时使数据块从节点进入传媒所需要的时间,也就是从数据块的第一个比特开始发送算起,到最后一个比特发送完毕所需的时间,也称为传输时延。传播时延是电磁波在信道中需要传播一定得距离而花费的时间。处理时延是指数据在交换节点为存储转发而进行的一些必要的处理所花费的时间,处理时延的长短取决于网络中当时的通信量。

IETF的IP性能度量工作组(IP Performance Metrics Working Group)进行网络性能度量的标准化工作,定义包括规定单向时延(One2way Delay)和往返时延(Round2trip Delay)。单向时延是指从发送方(源端)的物理设备开始发送数据至接收方(目的端)的物理设备全部接收到该测量报文所需要的时间。往返时延是指测量报文往返的时延测量,即测量报文从发送方开始发送时计时,接收方在接收到该报文后立即返回给发送方,发送方接收到该测量报文后停止计时,此时发送方所得的时间间隔即为往返时延。

5.4.2 时延测量中的关键技术和影响因素

在进行时延测量时首先要消除测量中出现的随机性。如同任何测量一样，对网络的时延测量存在很大的随机性。网络分组的时延是一个随时间变化的随机变量，由固定时延和可变时延两部分构成。固定时延是基本上不变的，它由传输时延和传播时延构成。传输时延由分组大小和链路容量决定，一个分组的大小一旦固定，通过的链路容量便是固定的，其传输时延也是固定的。传播时延由固定的物理传输介质确定并且是固定的。分组时延中的可变时延由很多因素造成，它可以分成中间路由器处理时延和排队等待时延两部分。对于任何一个分组，中间路由器总要对其进行路由查表以确定其转发端口，这个时间可以看作处理时间。同时，中间路由器繁忙可能导致分组排队等待处理，也需要一段等待时间。处理时间和等待时间是不固定的，由路由器的具体性能以及链路的拥塞状况而定，是一个随机变量。所以如果链路不出现拥塞，一个分组的最小时延便与该分组的大小成线性关系。分组的时延具有突发性和偶然性，为了能够使测量结果尽可能地反映网络的真实情况，可以采用低通滤波的方法来消除随机性。

其次，对网络时延的测量，必须保证网络入口点和出口点的时钟同步。如果时钟不同步，则时延测量会有很大的时钟误差。网络中路径往往是非对称的，或者即使路由是对称的但往返具有不同的性能特征。单向时延测量要求严格的时钟同步。GPS 时间同步技术是当前比较成熟的并在国际上广泛使用的时间同步技术。单向时延测量一般用在时间精度要求比较高的情况下，成本也比较高。这在实际的测量中，测量精度不是主要问题。网络时延通常在数十到数百毫秒范围内，测量存在几毫秒的误差完全可以接受。因此许多测量方案都采用往返时延，以避免时钟同步问题，本文中提到的时延均指往返时延。

目前，比较成熟的网络时延测量基本思想是，通过计算网络数据包的往返时延来评测网络的其他性能。往返时延是指发送方（源端）主机在 T_1 时刻将数据发往接收方（目的端）主机，接收方主机收到数据后马上把收到的数据发送到发送方主机，发送方主机在 T_2 时刻收到从接收方主机发回来数据的最后一个字节，则测试报文往返时延 R 就等于 $T_2 - T_1$ 。若发送端在指定的时间内没有收到从另一端发送的回应信息，就作为数据的一次丢包并记录下来。因为来回的时间是存在着差异的，且往返的路径有可能不同，所以 R 不一定能精确地反映出网络的单向时延。时延变化可以反映出带宽的变化，越大的时延说明了传输层带宽的利用率越低；最小值时延反映出了网络在最小负载情况下的性能，超过最小时延的部分，大致反映了网络拥塞的情况。在精度要求不是很高的情况下，可以用往返时延来估算单向时延，以及其他一些性能参数等。此外，通过周期性的长期测试可以反

映出这段时间里网络性能的变化。

此外,在采用主动测量时,难免受到“噪声”分组的影响。所谓“噪声”分组是指夹杂在探测分组当中,或处于探测分组之前、之后对测量结果造成影响的业务分组。例如,在采用分组对或多分组技术测量链路的瓶颈带宽时,难以保证探测分组在瓶颈链路处彼此相邻排队,可能会在中间插有其他分组,导致时间扩展(瓶颈带宽低估),或经过瓶颈链路后在第一个分组前插入其他分组导致时间压缩(瓶颈带宽高估)。使得时间压缩和时间扩展的分组就是测量过程中的“噪声”,消除这个噪声可采用以下方法。

① 求均值,但是由于“噪声”的随机性,该方法会造成较大误差;

② 在带宽估计的分布值中,选择密度最大的点,如采用直方图统计技术,但是事先不知道带宽的分布情况,直方图的条的宽度不好给出;

③ 采用在统计学中使用的非参数估计方法的核密度估计算法。

网络时延总是随着网络的运行状况而不断动态变化的,通过分析,可以把影响网络时延测量的因素归为以下两类。

(1) 网络本身的性能

网络本身的性能是指完全独立于网络测量方法、仅与网络本身特性相关的部分,主要包括:

① 网络的物理介质。

② 当前传输所需的路由、转发、过滤次数。

③ 网络的数据链路层协议的不同实现方式。

④ 当前网络负载情况。网络负载的大小直接影响网络时延测量结果,在轻负载情况下测出最小时延,超出最小时延的部分就在一定程度上反映了当前网络的负载状况。

(2) 测量报文的特定性能

测量报文的特定性能与网络具体的测量方式密切相关,主要可以分成:

① 承载测量报文的协议类型。

② 报文长度、发送频率。

③ 数据样本大小的选定。

此外,当前的网络流量、处理机性能和其当前的作业量等,也对网络时延产生影响。在实际测量中,要根据具体的测量环境来调整相应的测量策略。

要在不影响正常业务运营的情况下,测量网络时延才能达到对网络状态的真实反映。目前,基于端到端的网络时延测量方法,能够达到快速、准确测量的目的。这种基于端到端的时延测量是端到端之间的时延测量,不是单个设备的时延

测量，而是某一段甚至某一跳之间链路的时延测量。端到端往返时延计算的时间都是在同一台计算机上完成的，不存在时间的同步问题，无须在测量的两端都安装上测量软件，并且只需要固定一个测量点，而另一个测量点可以任意选择，从而使得测量的范围扩充到整个因特网。

端到端传输的优点是，链路建立后，发送端知道接收设备一定能收到，而且经过中间交换设备时不需要进行存储转发，因此时延较小。端到端传输的缺点是直到接收端收到数据为止，发送端的设备一直要参与传输。如果整个网络时延很长，那么对发送端的设备造成很大的浪费。端到端传输的另一个缺点是如果接收设备关机或故障，那么端到端传输不可能实现。在端到端网络时延测量中，要降低端到端测量的缺点带来的影响，就要选择合理的测量时段、数据包字长和测量样本等。

5.4.3 网络时延测量方法

根据网络时延测量其承载数据包的协议不同，网络时延测量方法可以分为以下几种。

(1) 基于 ICMP 协议的网络时延测量

Ping 是基于 ICMP 请求应答报文开发的应用软件，用来诊断网络故障，也是用来测量往返时延最常用的工具。Ping 工具通过向网络中特定地址发送 ICMP Echo 报文并监听回应报文，根据回应报文的探测结果，来评估主机路径的可靠性、往返时延大小等。将报文封装在 ICMP 报文中是目前网络测量工具最常用的方式，但是 ICMP 报文也是用来进行 DOS 攻击的主要方式，因此，许多 ICMP 协议都被过滤了，并有可能会完全被阻塞。路由器可以限制 ICMP 回复的速率来避免被消耗过多的网络带宽和路由资源，这意味着利用 ICMP 协议测量时延的结果是不可靠的。

(2) 基于 UDP 协议的网络时延测量

当对方的主机或者中间节点过滤掉 ICMP 报文后，可以使用 UDP 报文来实现网络的往返时延测量。UDP 报文的实现和 ICMP 报文较为相似，都是无连接的。不同的是 UDP 协议需要指定一个端口，在实际使用时，可以任选一个端口，当接收方该端口对应的服务不存在时，会立即返回一个“端口不可到达”的 ICMP 报文。因此，在使用 UDP 报文进行测量时，要注意发送方发送一个 UDP 报文，接收方反馈的不一定是 UDP 报文，也有可能是 ICMP 报文，这两种返回的报文都可以用来计算网络的时延。此外，UDP 协议在实现时要注意其最大报文长度与 ICMP 报文不同，通常应小于 500 Byte。但是对将报文封装在 UDP 报文中也有疑

问。UDP 没有 TCP 协议所固有的拥塞控制算法，在 UDP 报文使用达到高峰时，为了减少它们对 TCP 流的影响，UDP 报文的速率可能会受到限制。

(3) 基于 TCP 协议的网络时延测量

针对防火墙设置，其实现比较复杂，但可适用于测量某些前两种方法不能工作的特定环境。除了选取合适的承载报文外，TCP 协议测量网络时延的难点包括 TCP 协议发送机制本身。TCP 协议为了提高传输效率，往往不是即时发送报文，而要收集足够的数后才会发送，因此可能会产生发送方记录的发送时间远远早于时间发送时间，而使得测量的结果大于实际值的现象。为了避免这个现象的发生，要求在主动测量时，不但要求发送方在承载的 TCP 报文中加入 PSH 标志，此时将立即发送此数据报文而无须等待，同时也要求接收方在反馈该报文时，也必须带 PSH 标志。同理，在被动测量时，要注意承载的报文必须是带有 PSH 标志的报文，并且对应的反馈报文也必须同样带有 PSH 报文。

(4) 利用其他协议实现网络时延的测量

此方法基本和 UDP 或 TCP 类似，关键根据所在网络设置的具体情况，灵活选择承载报文，即能在发送方立即发送，而接收方收到后又能立即反馈的报文。

5.4.4 抖动测试方法

1. 浴缸曲线法（误码仪）

衡量系统功能的最终尺度是 BER。抖动分析的目标是确定抖动对 BER 的影响，并保证系统 BER 低于某个最大值，通常是 10^{-10} 。BERT 由码型发生器、误码分析仪和系统时钟组成。数据发生器把码型传送到系统器件上。器件处理码型，把结果传送到误码分析仪，误码分析仪在已知码型上同步，计算收到的码数，确定哪些码接收错误，计算 BER。

BERT 扫描或浴缸曲线如图 5.10 所示。

其中，误码率是样点时间位置的函数 $BER(t)$ 。

- Bit Error Ratio: 误码率。
- Eye Width at $BER = 10^{-10}$: $BER = 10^{-10}$ 时的眼宽。
- Time, t , relative to the reference clock: 相对于参考时钟的时间 t 。

图 5.10 是 BER 图，它是样点时间位置 $BER(t)$ 的函数，这个图称为 BERT 扫描图或浴缸曲线，简言之，它在相对于参考时钟给定的额定取样时间的不同时间 t 上测得的 BER。参考时钟可以是信号发射机时钟，也可以是从接收的信号中恢复的时钟，具体取决于测试的系统。

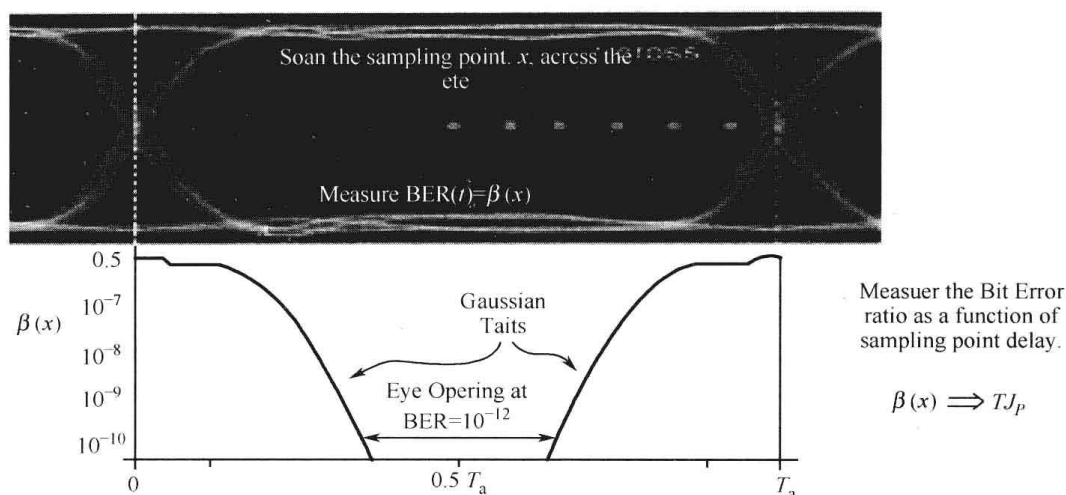


图 5.10 BERT 扫描或浴缸曲线

图 5.10 中的时间轴和眼图相同，两侧与眼图边沿相对应，样点位于中心。当 BER 一定时，曲线之间的距离是该 BER 上的眼图张开程度。在样点接近交点时，抖动会导致 BER 提高到最大 0.5。完整的 BERT 扫描测量是直接衡量 BER 一定时眼图张开程度的唯一方法。遗憾的是，完整的 BERT 扫描图需要很长的时间，如速率为 5 Gbps 时需要 30 分钟。通过集中在眼图边沿进行部分 BERT 扫描，可以在几秒内，在 10%~15% 的范围内保守地估算任何 BER 时的眼图张开程度。估算技术是一种近似方式，它对 RJ 和 DJ 概率分布去卷积，然后或多或少地使用高斯 RJ，从而计算 BER。关键在于，由于 DJ 有界，RJ 是高斯分布，卷积的分布尾部在 BER 下限中遵守高斯 RJ 分布。从高斯分布中推导出的函数，称为互补误差函数，拟合到 BERT 扫描平滑的下降沿上，然后可以推断拟合的函数，直到任何 BER 值，以估算眼宽。

在 BERT 扫描情况下，通过对 RJ 和 DJ 近似去卷积，可以从眼图交点直方图中估算 BER，基于这样一个事实，即在远离交点时，抖动分布主要取决于高斯 RJ。把直方图的尾部与高斯分布匹配起来，提供了一个函数，然后可以在眼图中推断这个函数，用来估算 BER 一定时的眼宽。数据集合和分析与高速取样时间间隔分析仪中基本相同。交点直方图拟合的宗旨与拟合 BERT 扫描的宗旨类似，但 BERT 扫描技术的速度和精度都要高得多。通过拟合直方图获得的某种高斯形状，在一定程度上受到随机波动的影响；BERT 扫描只取决于数据转换是否在样点上波动，而不取决于其波动的精确时间位置。因此，使用 BERT 进行测量要比低概率波动强健得多，后者的拟合可能会偏向交点直方图。部分 BERT 扫描的拟合速度要比交点直方图的拟合速度快得多，因为 BERT 数据集是在完整的数据速率上采集的，而直方图则是通过以低得多的速率对信号取样构建的。

对于高斯 RJ 能否精确地描绘交点直方图或 BERT 扫描的低 BER 的尾部，人们还存在争议。这种真正的随机流程会导致抖动，抖动遵守高斯分布并不存在问题，但这些流程可能并不是抖动分布尾部的主要因素。问题在于，多种小的效应

之和，会构成接近高斯的分布。统计理论的中心极限定理体现了这一点：数量无穷大的小流程之和会遵守高斯分布。这意味着多个低幅度的 DJ 流程会卷积到一个仅在钟形曲线中心、而不是尾部接近高斯分布的分布函数。但是尾部才是高斯假设对估算 BER 最重要的地方。可能在某些情况下，真正的 RJ 只占似乎是高斯分布的抖动的一小部分。如果是这样，那么用来从拟合快速 BERT 扫描测量和直方图拟合中估算 BER 所使用的技术，可能会在总体上高估实际 BER。尽管近似方法总是要出现错误的，但在高斯分布没有很好地描述尾部时，它可能会迫使制造商设计的抖动余量超过必要的水平。幸运的是，通过以 BER=10⁻¹² 的误码率水平执行全面的 BERT 扫描测量，进而可以把拟合技术与整个测量进行对比，检查这种情况并不难。

在当前的各种标准中，1GBE/10GBE 以太网设备、光纤通道（Fiber Channel）以及其他新的高速总线测试标准，应用误码仪采用浴缸曲线对系统进行误码率测试成为了抖动的最终评价标准。

2. 相噪分析法（相噪分析系统 / 信号源分析仪）

由于抖动也可以理解为相位噪声对时间的积分，因此测试数据信号的相位噪声也是抖动的一种方法。图 5.11 是时钟信号的相位频谱密度：每个单位的抖动-频率带宽的均方相位变化，它等于单边带（SSB）功率频谱。在图 5.11 中，RJ 提供了频谱中的连续背景，可以理解为闪烁、随机漂移和白噪声成分。通过分析相噪频谱，可以识别和分隔不同类型的 DJ：在图 5.11 中，在低抖动频率上，60 Hz 拾波及其谐波上升到连续背景上方的杂散信号，大约 2 kHz 上的宽拐角是锁相环滚降点，5 MHz 周围 PJ 非常明显。通过在希望的带宽上求积分，可以从频谱中提取 RMS 抖动。

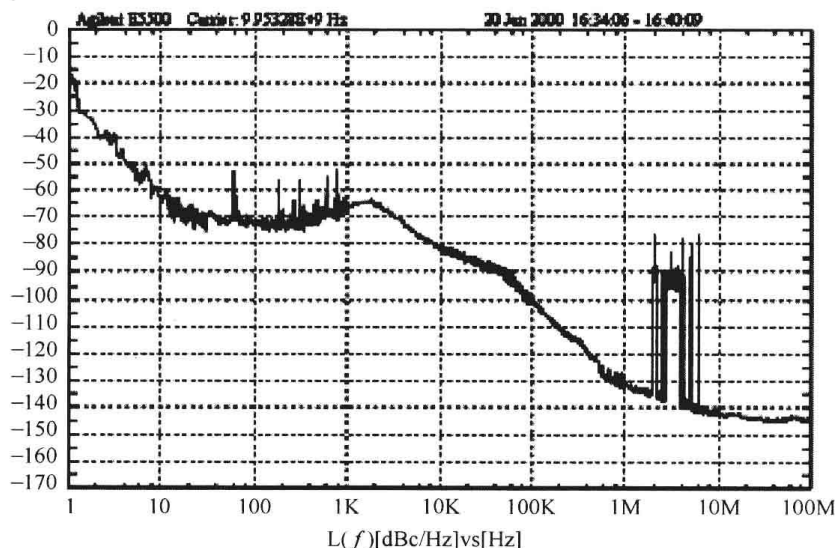


图 5.11 相噪频谱密度 / 单边带相噪图

通过使用具有专用相噪功能的频谱分析仪，还可以从时钟信号的频谱中提取SSB 频谱。与相位检测相比，安捷伦开放实验室暨测量方案中心的这种方法有两个系统问题。首选，频谱不能区分幅度噪声和相位噪声。低噪声、高带宽的限制放大器可以降低这种效应。其次，频谱分析仪的滤波形状允许载波附近的某些高幅度噪声泄漏出去。另外，使用频谱分析仪提取SSB 频谱的优点是带宽高。因此，结合使用抖动频率带宽高达约100 MHz 的相位检测器与具有覆盖最高时钟频率一半的相噪专用功能的频谱分析仪，可以提供一种强大的抖动诊断工具。

3. 鉴相法（传输分析仪）

分析抖动最直接的方法是通过鉴相法，即比较参考时钟和数据恢复时钟的时间间隔误差，从而解调出数据抖动随时间分布的曲线，最终通过标准规定的带通滤波器滤波，得到特定带宽的内的抖动幅度。

分析解调的抖动信号或相噪是一种强大的诊断技术。可以使用相位检测器解调相位。由于相位检测器只能解调时钟信号，要求专门的时钟恢复（CR）电路来分析数据信号。为把数据上的抖动传送到恢复的时钟中，而且没有失真，CR 必须有：

① 低抖动输出——CR 抖动提高了本底噪声，因为 CR 抖动的响应可能会干扰器件的抖动，而不能完全减去抖动响应。

② 平坦的转换——这样抖动信号不会失真，而且带宽要足够宽，能够在相关的抖动-频段中传送抖动。

4. 眼图测试法（示波器）

对于数字示波器而言，典型的抖动测试方法主要有以下两种。

① 采用数字存储示波器的等效采样模式或直接使用采样示波器，通过直方图统计测量累计定时抖动。等效采样的缺点是无法消除示波器自身的触发动抖对测试结果的影响，并且由于它采用的是多次触发、多次采集、累计显示的工作方式，对于电路设计和调试而言受到较多的限制，无法进行深层的抖动分析；另一个限制是该方法抖动测试参数有限，例如，不能测试周期间抖动。

② 更为流行的方法是采用数字存储示波器的实时捕获模式，单次触发、连续采集大量数据，配合相应的抖动测试软件进行抖动测试。当通过实时采集模式时，由于示波器工作在单次触发模式，连续实时采集所有信号，所以它不受仪器多次触发带来的触发动抖影响，并且它可以通过复杂的抖动分析和抖动分解得到每一个抖动分量，帮助设计和测试人员分析抖动产生的原因，甚至通过抖动分解估算系统的误码率。

5.4.5 Smartbits 丢包率测量方法

1. 连接 SmartBits 机箱并占用测试端口

单击菜单栏 Setup->SmartBits Connections, 在弹出窗口中添加 SmartBits 机箱的 IP 地址; 单击 OK 退出, 然后按 F8 键连接机箱, 如图 5.12 所示。

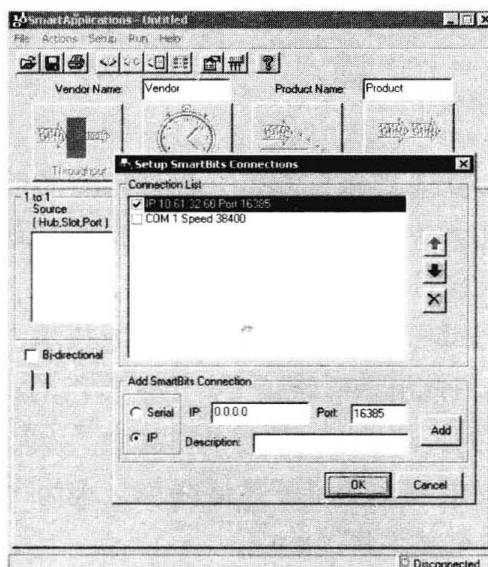


图 5.12 使用 Smartbits 进行测试 (1)

单击按钮栏的 Reserve/Release Smart Modules 按钮, 在弹出窗口中选择要使用的测试板卡, 然后单击右边的 Reserve 按钮占用端口。绿色表示未占用, 蓝色表示已占用, 如图 5.13 所示。

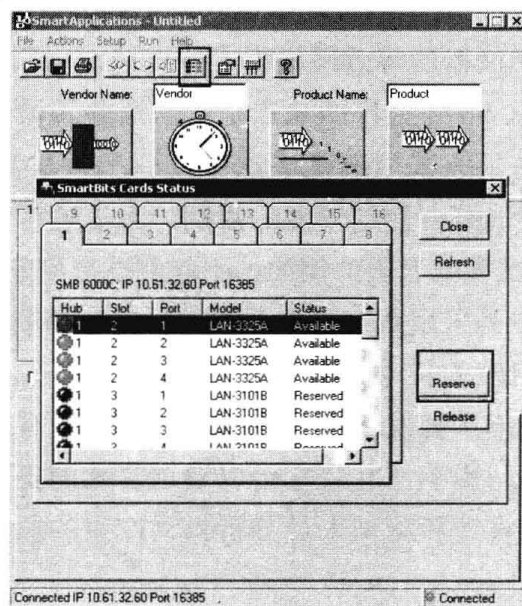


图 5.13 使用 Smartbits 进行测试 (2)

注意：默认情况下，连上机箱后，所有板卡都会被占用，请及时释放不用的板卡以便别人使用。

在端口列表中选择要使用的源端口和目的端口，然后用中间的 Add Selected Pair 按钮添加到右边的 Test Pairs 窗口，如图 5.14 所示。

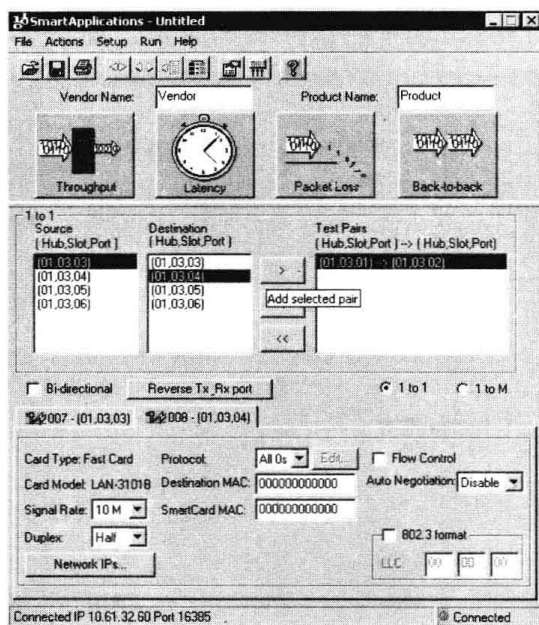


图 5.14 使用 Smartbits 进行测试 (3)

单击 Test Pair 窗口中的端口对，然后再下方各个端口的标签页中设置各个端口的物理属性和 MAC 地址。注意：端口的 Auto Negotiation 默认是 disable，需要改成 Force，如图 5.15 所示。

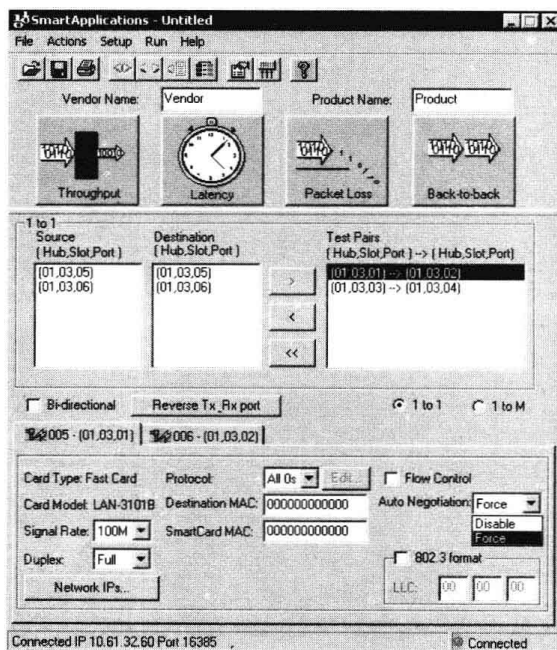


图 5.15 Smartbits 连接

2. 设置丢包率测试的参数

单击按钮栏中的 Setup Test Parameters 按钮，在弹出窗口中的右边选择 Packet Loss 标签页设置测试时长和次数。按 Use Custom 旁的 Sizes 按钮，在弹出窗口的下拉列表中选择 Packet Loss，设置测试帧长和时延，如图 5.16 所示。

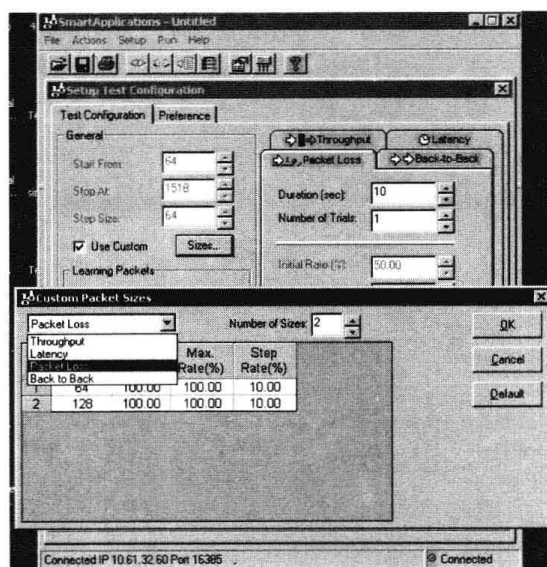


图 5.16 Smartbits 配置

3. 运行测试，查看和分析测试结果

单击 Run Packet Loss Test 按钮开始测试，如图 5.17 所示。

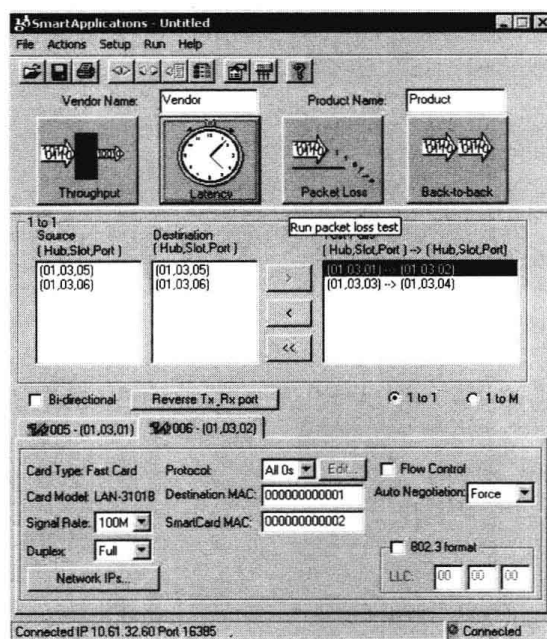


图 5.17 Smartbits 测试

测试结果如 5.18 所示。

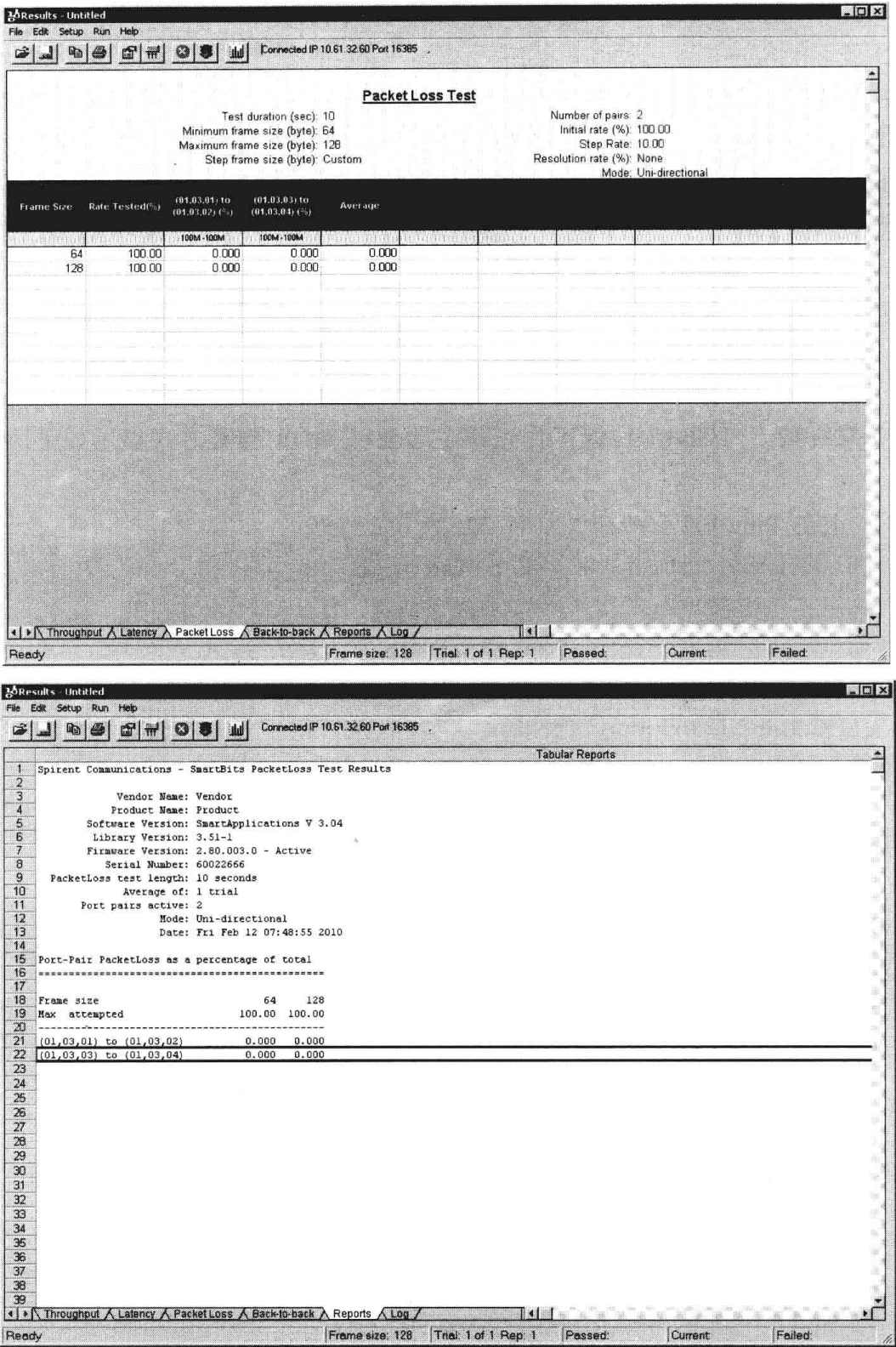


图 5.18 Smartbits 测试结果

最后在测试报告中给出端口至端口的丢包率。

5.4.6 利用 Ping 命令进行快速网络性能测试

我们知道可以用 Ping 命令来测试一下网络是否通畅，这在局域网的维护中经常用到，方法很简单，只需要在 DOS 或 Windows 的开始菜单下的“运行”子项中用 Ping 命令加上所要测试的目标计算机的 IP 地址或主机名即可（目标计算机要与你所运行 Ping 命令的计算机在同一网络或通过电话线或其他专线方式已连接成一个网络），其他参数都可不加。如要测试 IP 地址为 192.168.2.12 的终端与服务器是否已连网成功，就可以在服务器上运行 Ping 192.168.2.12 即可。

（1）TCP/IP 协议工作正常

如果终端上 TCP/IP 协议工作正常，即会以 DOS 屏幕方式显示如下所示的信息。

① 正在 Ping 192.168.2.12 具有 32 字节的数据。

来自 192.168.2.12 的回复: 字节=32 时间=6 ms TTL=121;

来自 192.168.2.12 的回复: 字节=32 时间=6 ms TTL=121;

来自 192.168.2.12 的回复: 字节=32 时间=5 ms TTL=121;

来自 192.168.2.12 的回复: 字节=32 时间=6 ms TTL=121。

② 192.168.2.12 的 Ping 统计信息。

数据包已发送 = 4;

数据包已接收 = 4;

数据包丢失 = 0 (0% 丢失)。

③ 往返行程的估计时间（以毫秒为单位）。

最短 = 5 ms;

最长 = 6 ms;

平均 = 5 ms。

从上面我们就可以看出目标计算机与服务器连接成功，TCP/IP 协议工作正常。

（2）网络未连成功

如果网络未连成功，则显示如下错误信息。

① 正在 Ping 192.168.2.12，具有 32 字节的数据。

请求超时;

请求超时;

请求超时;

请求超时。

② 192.168.2.12 的 Ping 统计信息。

数据包已发送 = 4;

数据包已接收 = 0;

数据包丢失 = 4 (100% 丢失)。

因为一般系统默认每次在用 Ping 测试时是发送 4 个数据包, 所以不管网络是否连通, 在提示信息中都会有出现 4 次反馈的信息。

如果我们要对一条对端 IP 已知且 ICMP 包不被路由设备拦截的链路进行快速的网络性能测试, 可以已使用 Ping 命令进行。

测试一定时间内的丢包率和网络时延平均值, 我们可以是用 -t 参数并配合计时器手动结束测试。-t 参数使 Ping 命令持续向对端地址发包, 直到手动强制停止。

(3) C:\Users\ALEX>ping -t 192.168.1.1

① 正在 Ping 192.168.1.1, 具有 32 字节的数据。

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=5 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=3 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间<1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=3 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=2 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=2 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间<1 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间=3 ms TTL=64;

来自 192.168.1.1 的回复: 字节=32 时间= 6ms TTL=64。

② 192.168.1.1 的 Ping 统计信息。

数据包已发送 = 18;

数据包已接收 = 18;

数据包丢失 = 0 (0% 丢失)。

③ 往返行程的估计时间 (以毫秒为单位)。

最短 = 0 ms;

最长 = 6 ms;

平均 = 1 ms。

(4) Control-C

^C

可见最后由于按下了 Control-C 键，强制终止了 Ping 命令的运行，而我们也得到了这段时间内的丢包率和平均往返延迟的数据。通过以上结果我们也可以看出，Ping 命令的最小时间颗粒是 1 ms，而在规模较小的网络或者直连专线的情况下，网络时延可能会小于 1 ms，这时不能使用 Ping 命令进行简单测试，而需要使用专业仪表进行相关的测试。

缩 略 语

AAC	Advanced Audio Coding	高级音频编码
AC	Arithmetic Coding	算术编码
ACF	Admission Confirm	接入确认
ACK	Acknowledgement Message	应答消息
ACR	Absolutely Category Rate	绝对尺度分级法
AES	Audio Engineering Society	音频工程师协会
AG	Access Gateway	接入网关
AMG	Access Media Gateway	接入媒体网关
API	Application Programming Interface	应用程序接口
ARQ	Admission Request Message	接入请求消息
ATSC	Advanced Television Systems Committee	先进电视系统委员会
AVC	Advanced Video Coding	先进视频编码
AVS	Audio Video Coding Standard	中国音视频编码标准
AWGN	Additive White Gaussian Noise	加性高斯白噪声
B 帧	Bidirectionally Predicted Picture	双向预测编码图像帧
BER	Bit Error Ratio	误码率
BET	Blur Edge Time	模糊边沿时间
BGCF	Breakout Control Function	出口网关控制功能
BICCP	Bearer Independent Call Control Protocol	与承载无关的呼叫控制协议
CIE	International Commission on Illumination	国际照明委员会
CIF	Common Intermediate Format	通用中间格式
CMMB	China Mobile Multimedia Broadcasting	中国移动多媒体广播
CN	Core Network	核心网
CRC	Cyclic Redundancy Code	循环冗余校验
CSCF	Call Session Control Function	多媒体呼叫会话过程中的信令控制
CVBS	Composite Video Broadcast Signal	复合视频信号
DAB	Digital Audio Broadcasting	数字音频广播

DCF	Disengage Confirm	退出确认
DCT	Discrete Cosine Transform	离散余弦变换
DFT	Discrete Fourier Transform	离散傅里叶变换
DHCP	Dynamic Host Configuration Protocol	动态主机配置协议
DIP	Direct Intra Prediction	直接帧内预测
DMB	Digital Multimedia Broadcasting	数字多媒体广播
DMOS	Differential Mean Opinion Score	差异主观评价分
DNS	Domain Name Server	域名服务器
DRQ	Disengage Request	退出请求
DSCQS	Double-stimulus Continuous Quality Scale	双激励连续质量分级法
DTMB	Digital Terrestrial Multimedia Broadcast	数字地面多媒体广播
DVB	Digital Video Broadcasting	数字视频广播
DVI	Digital Visual Interface	数字视频接口
ES	Elementary Stream	基本流
ETSI	European Telecommunication Standards Institute	欧洲电信标准协会
EUT	Equipment Under Test	被测设备
EMC	Electromagnetic Compatibility	电磁兼容
FDM	Frequency Division Multiplexing	频分复用
FEC	Forward Error Correction	前向纠错
FFT	Fast Fourier Transform	快速傅里叶变换
FIFO	First In First out	先进先出
GK	Gate Keeper	网闸
GOP	Group of Picture	图像组
GUI	Graphical User Interface	图形用户接口
GW	Gateway	网关
HA	Home Agent	归属代理
HDMI	High Definition Multimedia Interface	高清晰度多媒体接口
HDTV	High Definition TV	高清晰度电视
HVS	Human Visual System	人类视觉系统
I 帧	Intra Coded Picture	帧内编码图像帧
IAD	Integrated Access Device	综合接入设备
I-CSCF	Interrogating CSCF	询问 CSCF
ICT	Integer Cosine Transform	整数余弦变换
IEC	International Electrotechnical Commission	国际电工委员会
IEEE	Institute of Electrical and Electronics Engineers	电气和电子工程师学会

IETF	Internet Engineering Task Force,	互联网工程任务组
IFFT	Inverse Fast Fourier Transformation	反向快速傅里叶变换
IM	Instant Message	即时消息
IM-MGW	IP Multimedia Subsystem Media Gateway Function	IP 多媒体网关功能
IMS	IP Multimedias Subsystem,	IP 多媒体系统
IN	Intelligent Network	智能网
ISC	IMS Service Control	IMS 业务控制
ISDN	Integrated Services Digital Network	综合业务数字网
ISO	International Organization for Standardization	国际标准化组织
ITU	International Telecommunication Union	国际电信联盟
JNE	Just Noticeable Difference	最小可觉差
JPEG	Joint Photographic Experts Group	联合图片专家组
JVT	Joint Video Team	联合视频专家组
LCD	Liquid Crystal Display	液晶显示器
LED	Light Emitting Diode	发光二极管
MC	Multipoint Controller	多点控制器
MCU	Multipoint Control Unit	多点控制单元
MDSP	Multimedia Message Service	多媒体消息业务
MGC	Media Gateway Controller	媒体网关控制器
MGCF	Media Gateway Control Function	媒体网关控制功能
MGCP	Media Gateway Control Protocol	媒体网关控制协议
MGW	Media Gateway	媒体网关
MOS	Mean Opinion Score	平均意见分数
MP	Multipoint Processor	多点处理器
MPEG	Moving Pictures Experts Group	动态图像专家组
MPRT	Motion Picture Response Time	运动图像相应时间
MTBF	Mean Time Between Failure	平均失效间隔时间
MRF	Media Resource Function	媒体资源功能
MRFC	Media Resource Function Controller	媒体资源功能控制
MRFP	Media Resource Function Processor	媒体资源功能处理
NBS	National Bureau of Standards	美国国家标准局
NRZ	Non Return to Zero	非归零
NTP	Network Time Protocol	网络时间协议
NTSC	National Television Standards Committee	国家电视标准委员会
OLED	Organic Light Emitting Display	有机发光显示器
P 帧	Predictive Coded Pictute	预测编码图像帧

PAL	Phase Alteration Line	逐行倒相
PBN	Packet Based Network	分组通信网络
PCF	Package Control Function	分组控制功能
PCM	Pulse Code Modulation	脉冲编码调制
P-CSCF	Proxy CSCF	代理呼叫会话控制功能
PDCP	Packet Data Convergence Protocol	分组数据汇聚协议
PDP	Plasma Display Panel	等离子显示器
PQR	Picture Quality Rating	图像质量评价
PS	Program Stream	节目流
PSNR	Peak Signal-to-Noise Ratio	峰值信噪比
QAM	Quadrature Amplitude Modulation	正交幅度调制
QCIF	Quarter Common Intermediate Format	标准化图像格式 176×144 像素
QoE	Quality of Experience	用户体验
QOS	Quality of Service	服务质量
QPSK	Quaternary Phase Shift Keying	四相相移键控
RADIUS	Remote Authentication Dial-in User Service	远端拨入用户验证服务
RLE	Run Length Encoding	行程编码
ROI	Region of Interest	兴趣区域
RSVP	Resource Reservation Protocol	资源预留协议
RTCP	Real time Transfer Control Protocol	实时传输控制协议
RTP	Real time Transfer Protocol	实时传输协议
RTSP	Real time Transfer Session Protocol	实时传输会话协议
SAG	Service Access Gateway	业务接入网关
SAP	Service Access Part	业务接入部分
SCP	Service Control Point	业务控制点
SCF	Service Capability Feature	业务能力特征
SCP	Service Control Point	业务控制点
SCS	Service Capability Server	业务能力服务器
S-CSCF	Service CSCF	服务 CSCF
SCTP	Stream Control Transmission Protocol	流控制传输协议
SDI	Serial Digital Interface	数字分量串行接口
SDP	Session Description Protocol	会话描述协议
SDTV	Standard Definition TV	标准清晰度电视
SDU	Service Data Unit	业务数据单元
SG	Signaling Gateway	信令网关
SIGTRAN	Signaling Transport	信令传递

SI	Service Information	业务信息
SIP	Session Initiation Protocol	会话初始化协议
SNMP	Simple Network Management Protocol	简单网络管理协议
SynC	Synchronization Channel	同步信道
TDM	Time Division Multiplexing	时分复用
TS	Transport Stream	传输流
UA	User Agent	用户代理
UAC	User Agent Client	用户代理客户端
UAS	User Agent Server	用户代理服务器
UDP	User Datagram Protocol	用户数据包协议
VIG	Video Interworking Gateway	视频互通网关
VGA	Video Graphics Array	视频图形阵列
VLC	Variable Length Coding	变长编码

参 考 文 献

- [1] 国家广播电视产品质量监督检验中心 TCL 多媒体科技控股有限公司 GPC 研发中心. 数字电视与平板电视中的色度学. 北京: 人民邮电出版社, 2010.
- [2] GY/T 241-2009. 高清晰度有线数字电视机顶盒技术要求和测量方法.
- [3] YD/T 1607-2007. 数字移动终端图像及视频传输特性技术要求和测试方法.
- [4] 高文, 赵德斌, 马思伟. 数字视频编码技术原理. 北京: 科学出版社, 2010.
- [5] 国家广播电视产品质量监督检验中心, 中国电子科技集团公司第三研究所, 天津大学电子信息工程学院. 数字电视原理与检测技术. 北京: 人民邮电出版社, 2011.
- [6] 刘东明, 等. 移动通信增值业务技术详解. 北京: 人民邮电出版社, 2009.
- [7] 梅玉平, 等. 3G 的业务及管理. 北京: 人民邮电出版社, 2007.
- [8] 西刹子. 智能网络视频监控技术详解与实践. 北京: 清华大学出版社, 2010.
- [9] 赵坚勇. 电视原理与系统 (第二版). 西安: 西安电子科技大学出版社, 2003.
- [10] 梅玉平, 等. 3G 与固定视频业务的融合. 北京: 人民邮电出版社, 2005.
- [11] 叶朝阳. 固定于移动融合 (FMC) 技术. 北京: 人民邮电出版社, 2008.
- [12] 高伟. 基于 SIP 的 IP 电话研究与实现. 西南交通大学, 2007.4.
- [13] 姚琴琴. 基于 SIP 协议的 IP 组播桌面视频会议系统的研究与设计. 苏州大学, 2007.4.
- [14] 陈萌. SIP 协议内容及其代理服务器的设计与实现. 重庆科技学院学报 (自然科学版), 2007.
- [15] 周海华, 边恩炯. SIP 原理与应用. 北京: 机械工业出版社, 2006.
- [16] 罗海文. 基于 IMS 的固定与移动网络融合. 北京邮电大学, 2006.
- [17] 杨鹤男. 基于 SIP 协议的视频会议系统的研究与实现. 北京邮电大学, 2007.
- [18] 蔚晓丹. 交通信息网与电信网的融合及 ITS 评价指标体系研究. 北京邮电大学, 2010.
- [19] 荀兆勇, 廖建新, 朱晓民. IMS 网络中业务触发数据的研究. 高技术通讯, 2008.12.
- [20] 蔡兵, 韩妮, 徐展琦. 一种基于 IMS 的呼叫中心实现方案. 电信工程技术与标准化, 2007.
- [21] 孙小红. 企业多媒体业务信息化 (视频会议) 应用研究. 天津市电视技术研究会, 2012 年年会论文集.
- [22] 冯传奋. 基于 IMS 的网络融合方案及演进探讨. 中国通信学会信息通信网络技术委员会 2009 年年会论文集 (下册).
- [23] 潘艺鹏. 构建基于网络视频会议的远程互动教学系统. 信息技术环境构建与教学应用, 2009.
- [24] 戚琦. IMS 会话控制关键技术研究. 北京邮电大学, 2010.
- [25] 程银波. 分布式 IP 多媒体业务网络架构及关键技术研究. 北京邮电大学, 2010.

Images have been losslessly embedded. Information about the original file can be found in PDF attachments. Some stats (more in the PDF attachments):

```
{
  "filename": "6KeG6aKR5Lya6K6u57O757uf5Y6f55CG5LiO5rWL6K+VXzEzNDYxMjQxLnV2eg==",
  "filename_decoded": "\u89c6\u9891\u4f1a\u8bae\u7cfb\u7edfu539fu7406\u4e0e\u6d4b\u8bd5_13461241.uvz",
  "filesize": 49032540,
  "md5": "1e2839cbcd1077715db87e435808bdf9",
  "header_md5": "9831f5bf49bfc56ce3310c73778563eb",
  "sha1": "761d1ed244f5b220de53ffffd70fb540bba5a6bb",
  "sha256": "0578fec92aba27e0f590a82cae7a6b41fc2acbce78f0709cc6c0118abda3b90c",
  "crc32": 666104254,
  "zip_password": "",
  "uncompressed_size": 56419112,
  "pdg_dir_name": "\u2569\u2559\u255e\u2561\u2557\u00df\u2565\u0398\u2567\u2561\u2550\u2502\u2558\u00a1\u2514\u03c6\u2559\u03b4\u2593\u0393\u2569\u2558_13461241",
  "pdg_main_pages_found": 182,
  "pdg_main_pages_max": 182,
  "total_pages": 192,
  "total_pixels": 1254654160,
  "pdf_generation_missing_pages": false
}
```