

YD

中华人民共和国通信行业标准

YD/T 3244—2017

远程呈现视频会议系统 系统架构

Telepresence conference system
system architecture

2017-04-12 发布

2017-07-01 实施

中华人民共和国工业和信息化部 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
3.1 缩略语	1
3.2 术语和定义	1
4 远程呈现视频会议系统架构	2
4.1 系统总体架构	2
4.2 点对点通信	3
4.3 多点通信会议	4
5 远程呈现视频会议系统组件	4
5.1 组件构成	4
5.2 远程呈现视频会议终端	6
5.3 MCU	7
5.4 CRC	7
5.5 GW	7
5.6 管理系统	7
5.7 传统视频会议终端	8
6 远程呈现视频会议系统呼叫流程	8
6.1 概述	8
6.2 H.323 协议体系	8
6.3 SIP 体系	12
6.4 H.323 与 SIP 互通	15

前 言

本标准是“远程呈现视频会议系统”系列标准之一，该系列标准预计的结构及名称如下：

- 《远程呈现视频会议系统业务需求》；
- 《远程呈现视频会议系统系统架构》；
- 《远程呈现视频会议系统协议技术要求》；
- 《远程呈现视频会议系统服务质量要求》。

本标准按照 GB/T 1.1—2009 给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由通信标准化协会提出并归口。

本标准起草单位：中国信息通信研究院，华为技术有限公司，中兴通讯股份有限公司，中国移动通信集团公司。

本部分主要起草人：孙明俊、肖 晶、叶小阳、王 亮、栗 蔚、郑建平、陈 曦。

远程呈现视频会议系统

系统架构

1 范围

本标准确定了远程呈现视频会议系统协议技术要求中的系统架构,主要包括远程呈现视频会议系统的整体架构、场景、系统组件、子系统架构以及总体呼叫流程。

本标准适用于远程呈现视频会议系统。...

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅所注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

[IETF CLUE FRAMEWORK] 远程呈现系统多码流系统框架(IETF ControLling mUltiple streams for tElepresence Framework)

3 术语、定义和缩略语

3.1 缩略语

下列缩略语适用于本文件:

CRC	呼叫资源控制器	Call and resource controller
GK	网守	Gatekeeper
GW	网关	Gateway
ISDN	综合业务数字网	Integrated Services Digital Network
MCU	多点控制单元	Multipoint Control Unit
PC	个人电脑	Personal computer
QoS	业务质量	Quality of Service
RAS	注册、准入和状态	Registration, Admission and Status
RTP	实时传输协议	Real-time Transport Protocol
RTCP	RTP 控制协议	RTP Control Protocol
SIP	会话初始协议	Session initiation protocol

3.2 术语和定义

下列术语和定义适用于本文件:

3.2.1

远程呈现 Telepresence

一种交互式音视频通信方式，通过高质量的音视频、眼神交流、注视感知、身体语言、空间音频、真人大小图像显示，以及协调的环境等一系列技术，使用户获得和面对面交流类似的体验和感受。

3.2.2

远程呈现视频会议系统 Telepresence Conference System

一系列功能、设备和网络元素的组合，在远程呈现视频会议中，对高质量的交互式音视频信号进行捕获、传输、管理和渲染。通过适当数量的设备（例如摄像头，屏幕，麦克风，扬声器，编解码器，MCU，PC）和特定的环境特征来达到远程呈现的体验。

3.2.3

渲染 Render

从媒体生成表示的过程，例如显示动态视频，或从扬声器发出声音。

3.2.4

音频混合 Audio Mixing

将多个音频信号标记累加，并产生一个单独的音频流。

4 远程呈现视频会议系统架构

4.1 系统总体架构

一个完整的远程呈现视频会议系统总体架构包括：远程呈现视频会议终端，传统视频会议终端，传统的音频终端，MCU，CRC，GW，管理系统等。视频呈现视频会议终端的样式可以有多种，可以是三屏、双屏或单屏等样式。远程呈现视频会议系统的架构既需要支持同构远程呈现视频会议系统之间的互操作性，也要支持异构的互操作性，如远程呈现系统和传统的视频会议系统以及音频会议。在系统架构中，可以有多个网关，分别连接到不同类型的网络环境，诸如固网和移动网络。系统总体架构如图 1 所示。

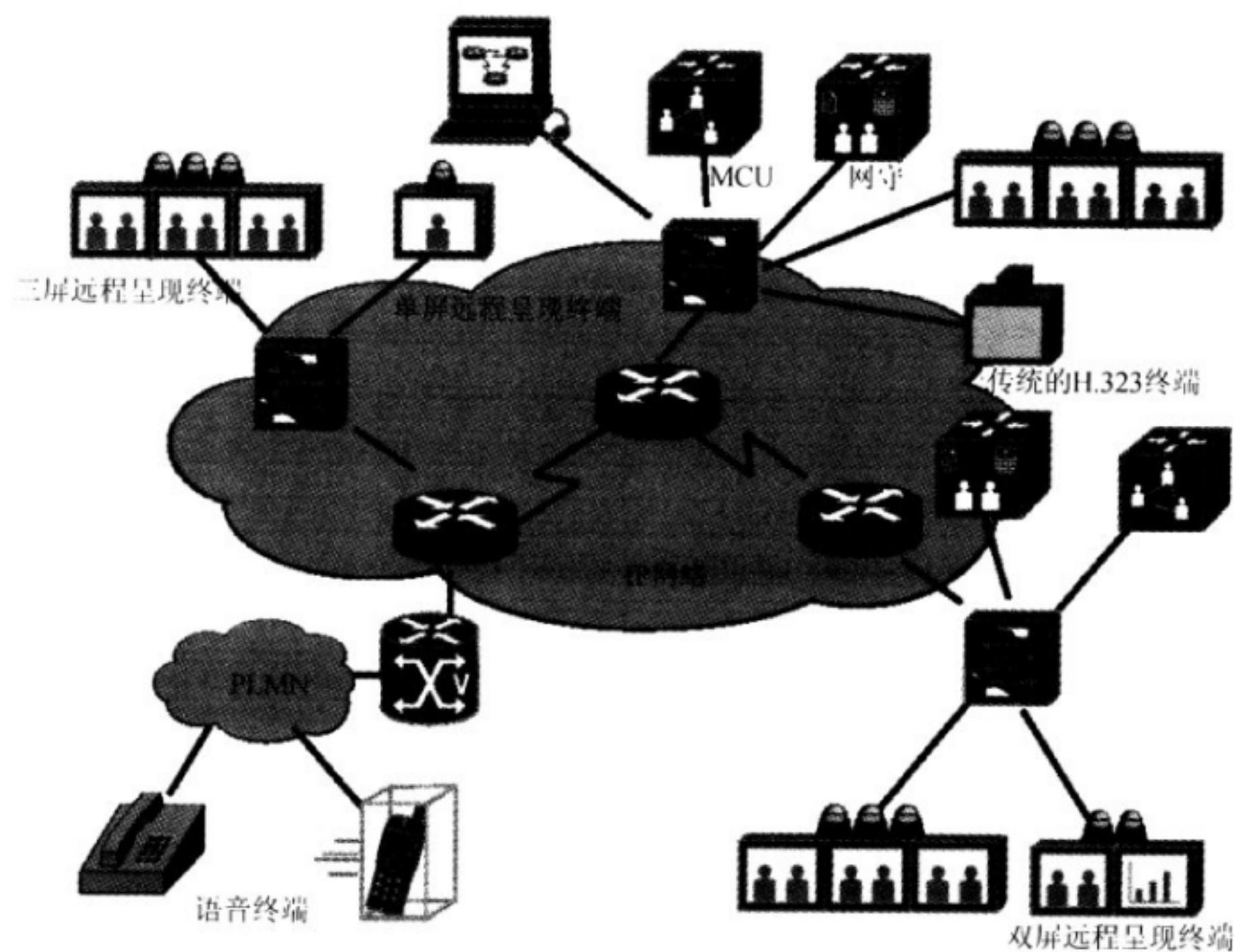


图 1 远程呈现视频会议系统总体架构

4.2 点对点通信

典型的远程呈现视频会议系统的点对点通信场景如图 2 所示。

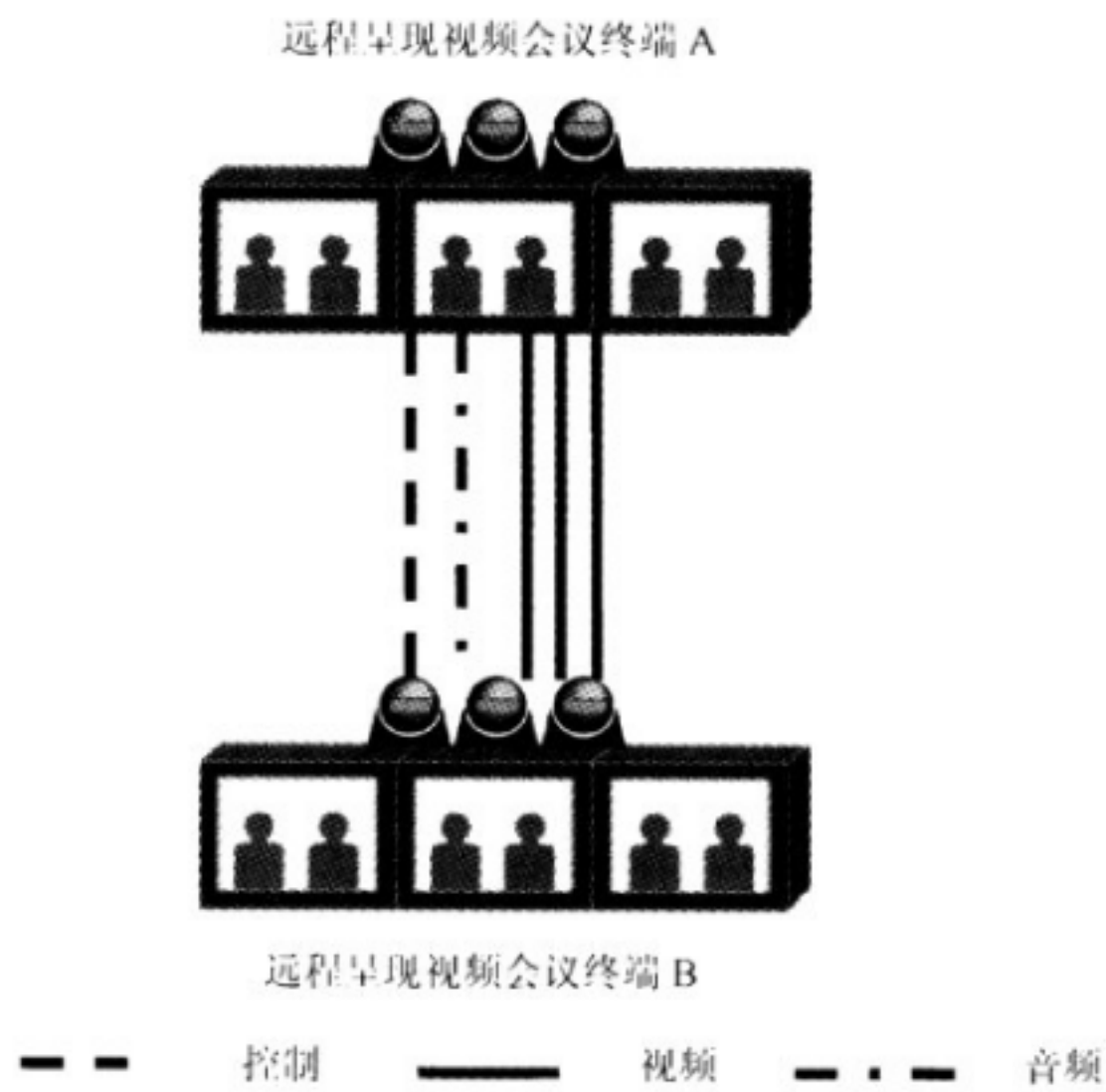


图 2 远程呈现视频会议系统点对点通信场景

图 2 中：

- a) 远程呈现视频会场之间总体采用 1 路控制，1 路音频码流（为多声道音频码流），以及多路视频码流的方式；
- b) 控制链路用于控制整个呼叫过程，如：启用/停止辅流发送等。

4.3 多点通信会议

典型的远程呈现视频会议系统的多点通信场景如图 3 所示。

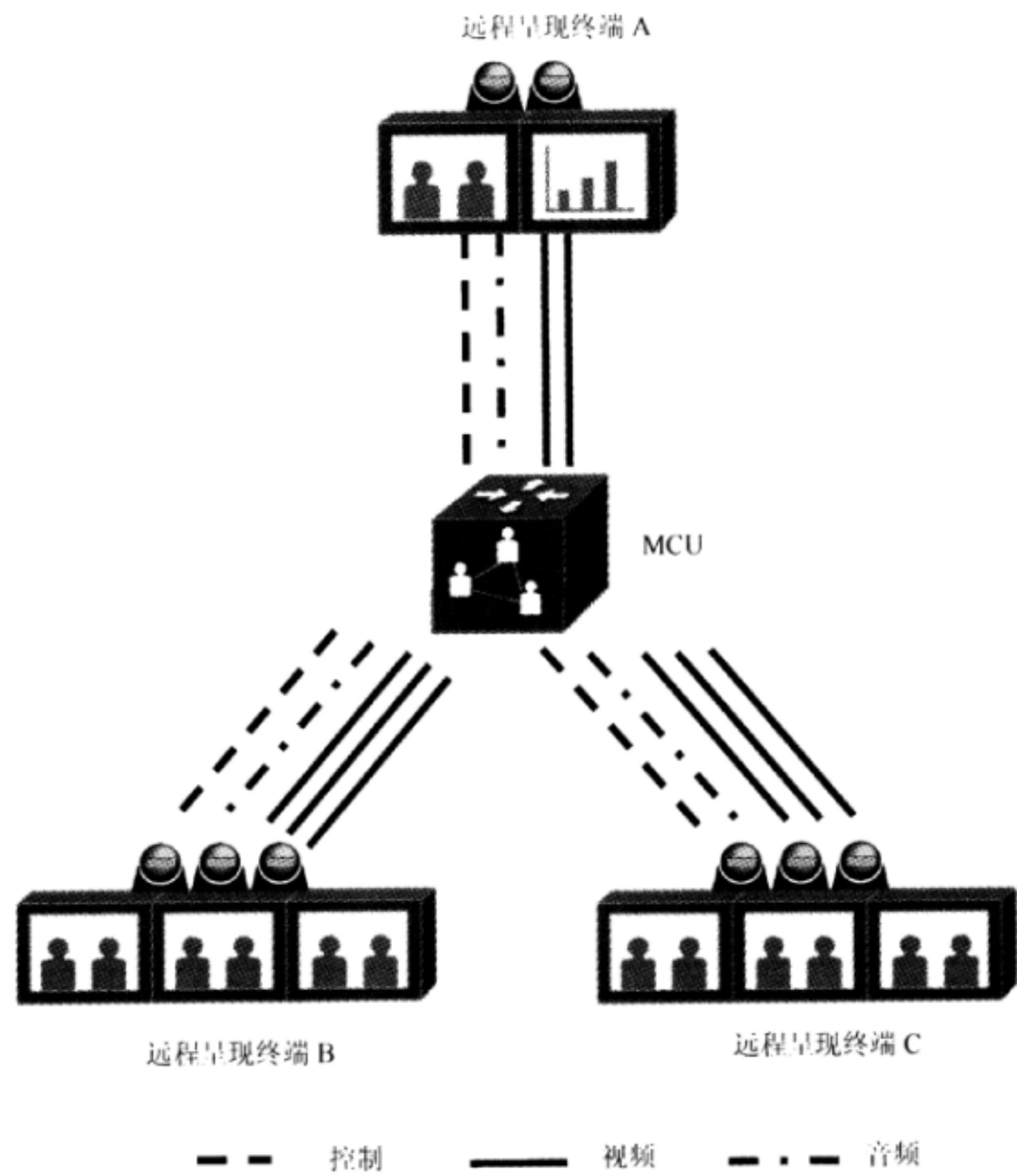


图 3 远程呈现视频会议系统多点通信场景

图 3 中：

- a) 不同会场之间通过 MCU 进行连接；
- b) 各终端根据其与 MCU 能力协商的结果，发送对应的码流；
- c) 由 MCU 完成对媒体的适配处理。如：音频单声道和多声道之间的适配、多路视频媒体之间的转发处理、会议控制的适配处理。

5 远程呈现视频会议系统组件

5.1 组件构成

远程呈现视频会议系统由数个功能组件构成，如图 4 所示。

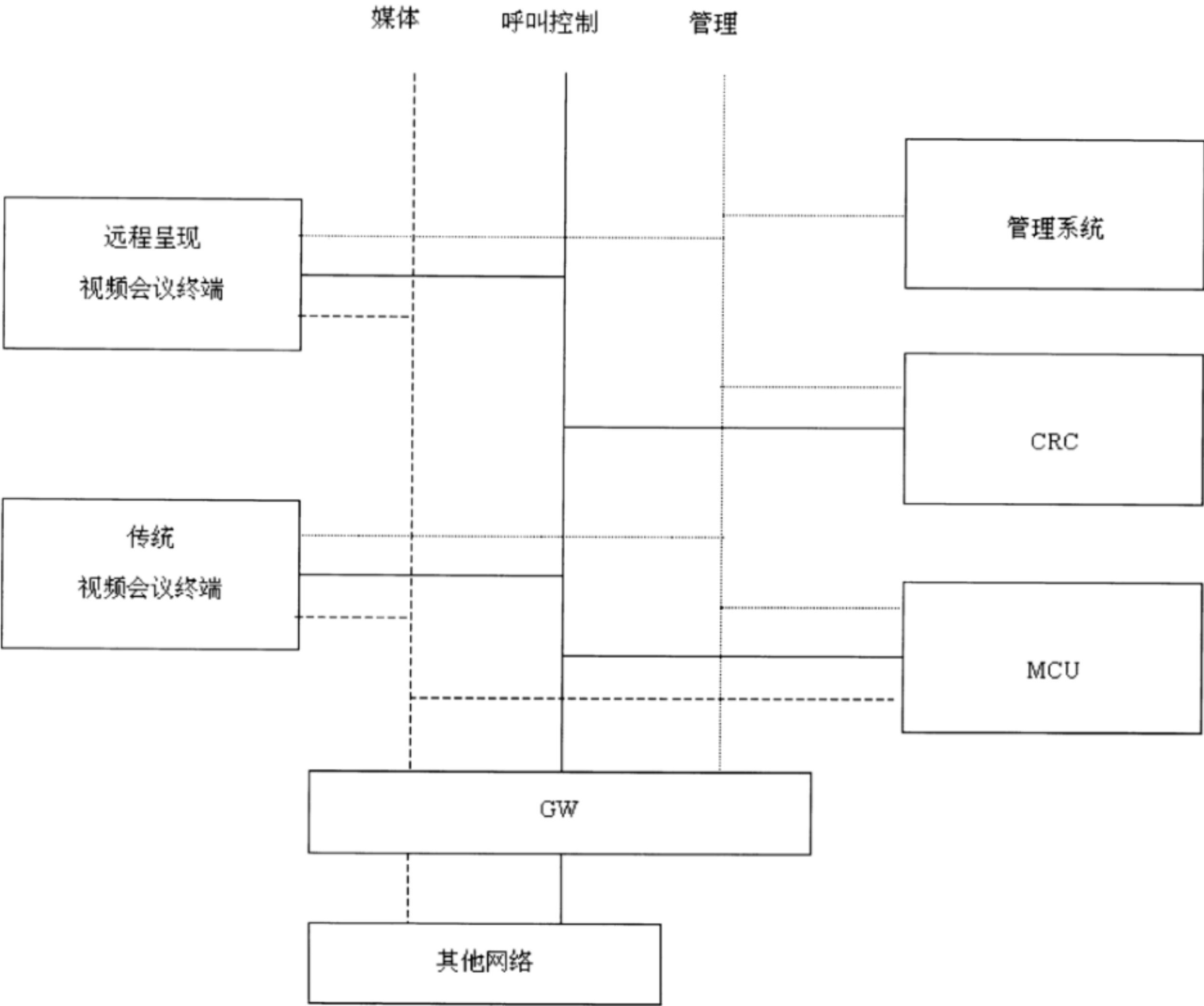


图 4 远程呈现视频会议系统功能组件

远程呈现视频会议系统功能组件，以及它们之间的媒体和呼叫控制的接口。所有的接口都是双向接口。另外，这些组件都可能有多实例，因此也会有多个媒体和呼叫控制的接口。这些接口分别在表 1、表 2 和表 3 中列出。

表 1 中列出的是呼叫控制接口。

表 1 远程呈现视频会议系统呼叫控制接口

编号	组件	
C1	远程呈现视频会议终端	CRC
C2	远程呈现视频会议终端	MCU
C3	远程呈现视频会议终端	GW
C4	远程呈现视频会议终端	传统视频会议终端
C5	远程呈现视频会议终端	远程呈现视频会议终端
C6	传统视频会议终端	CRC
C7	传统视频会议终端	MCU

表 1 远程呈现视频会议系统呼叫控制接口（续）

编号	组件	
C8	传统视频会议终端	GW
C9	传统视频会议终端	传统视频会议终端
C10	CRC	MCU
C11	CRC	GW
C12	CRC	CRC
C13	MCU	GW
C14	MCU	MCU

表 2 中列出的是管理接口：

表 2 远程呈现视频会议系统管理接口

编号	组件	
MT1	CRC	管理系统
MT2	MCU	管理系统
MT3	GW	管理系统
MT4	远程呈现视频会议终端	管理系统
MT5	传统视频会议终端	管理系统

表 3 中列出的是媒体接口：

表 3 远程呈现视频会议系统媒体接口

编号	组件	
M1	远程呈现视频会议终端	MCU
M2	远程呈现视频会议终端	远程呈现视频会议终端
M3	远程呈现视频会议终端	GW
M4	远程呈现视频会议终端	传统视频会议终端
M5	传统视频会议终端	传统视频会议终端
M6	传统视频会议终端	GW
M7	传统视频会议终端	MCU
M8	MCU	GW
M9	MCU	MCU

5.2 远程呈现视频会议终端

远程呈现视频会议终端是逻辑上的终结节点，它可以包含多个物理设备。远程呈现视频会议终端的主要功能是：

- a) 通过能力协商建立呼叫会话；
- b) 将本地媒体流捕获、编码、封包并传输；
- c) 将远端媒体流接收、拆包、解码并渲染。

与远程呈现视频会议终端不同，MCU 也可以发送和接收媒体流，但它并不会通过捕获和/或渲染来

发起或终结媒体流。远程呈现视频会议终端的形式多种多样：可以是3屏终端，单屏终端，或者音频终端等。

远程呈现视频会议终端的特征包含：

- a) 主流摄像头的数量、摄像头空间关系、捕获能力（比如分辨率，帧率和纵横比等）；主流摄像头的视点和视场，包括重叠区域、盲区，以及能否在所有深度范围内拼接无错位；
- b) 辅流摄像头的数量、捕获能力（比如分辨率，帧率和纵横比等）；
- c) 用于数据共享的视频流能力，比如分辨率，帧率和纵横比等；
- d) 主流屏的数量、类型、尺寸，主流屏的空间关系和能力（比如分辨率，帧率，纵横比和边框宽度等）；
- e) 辅流屏的数量、类型，辅流屏的空间关系和能力（比如分辨率，帧率，纵横比和边框宽度等）；
- f) 麦克风的音频信号的空间关系和麦克风的捕获能力；
- g) 扬声器的空间关系和扬声器的播放能力；
- h) 多点会议的显示策略，比如摄像头到显示器之间是静态映射还是动态映射等。

总的来说，远程呈现视频会议终端的特征为：在会话初始阶段通过信令流传输，用于在不同远程呈现视频会议终端之间进行能力协商。

5.3 MCU

MCU 是将两个或多个远程呈现视频会议终端连接到一个远程呈现视频会议的设备。MCU 可以建立并控制多个视频会议呼叫会话。在从多个终端或级联 MCU 和 GW 接收到媒体流后，MCU 处理这些媒体流，最终将处理后的媒体流发送到终端、MCU 或 GW。

MCU 应支持音视频混合以及不同类型终端的媒体适配，还应支持控制功能。比如辅流令牌控制、基于终端屏幕的声控切换和控制等。

5.4 CRC

CRC 是用于视频会议的资源管理设备，可以为终端、MCU 和 GW 提供资源管理。CRC 的功能包括：地址解析、接入控制和带宽管理。在 H.323 网络中，CRC 通常由 GK 实现；在 SIP 网络中，CRC 则由注册服务器实现。

5.5 GW

GW 可连接不同类型的网络（比如 H.323 和 SIP），并提供信令和媒体流的互操作。GW 也可以对具有不同特性（例如编码格式）的媒体流进行适配。GW 可以使位于不同网络（例如运营商网内网间，运营商与企业网之间等）的远程呈现视频会议终端、传统视频会议终端和其它组件都能接入同一个会议。

5.6 管理系统

管理系统主要包含控制和管理功能，以提供远程呈现视频会议服务。例如设备管理、会议预定和用户鉴权等。管理系统还可以提供用于远程呈现视频会议会话和系统使用相关的统计信息和报告等功能。

5.7 传统视频会议终端

传统视频会议终端分为 H.323 视频会议终端和 SIP 视频会议终端。H.323 视频会议终端的主要能力和功能应按照 ITU-T H.323 中的定义。SIP 视频会议终端可以是 SIP 硬终端和 SIP 软终端（如 PC 或笔记本电脑）。SIP 视频会议终端可通过 GW 连接到其它组件（例如 CRC）。

6 远程呈现视频会议系统呼叫流程

6.1 概述

对远程呈现视频会议系统来说，目前主流的有两大协议族：H.323 和 SIP，因此下面也以这两个不同的协议为基础，分别描述远程呈现视频会议系统的整个呼叫流程。

6.2 H.323 协议体系

6.2.1 概述

在整体上，远程呈现视频会议系统仍然基于现有 H.323 协议基本框架，如图 5 所示。

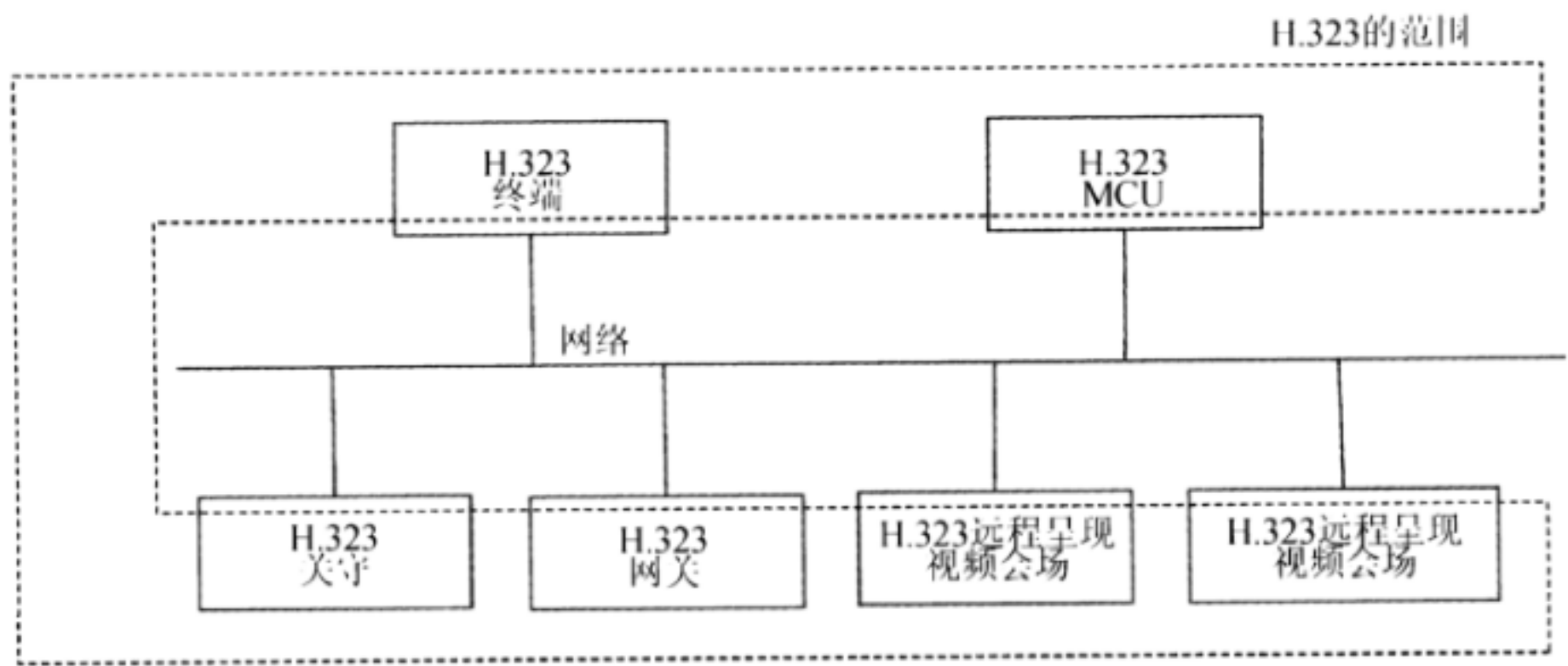


图 5 H.323 远程呈现网络模型

其中远程呈现视频会场作为 H.323 终端中的一种。这样，对远程呈现视频会议系统，其涉及的协议/控制信令的内容如图 6 所示。

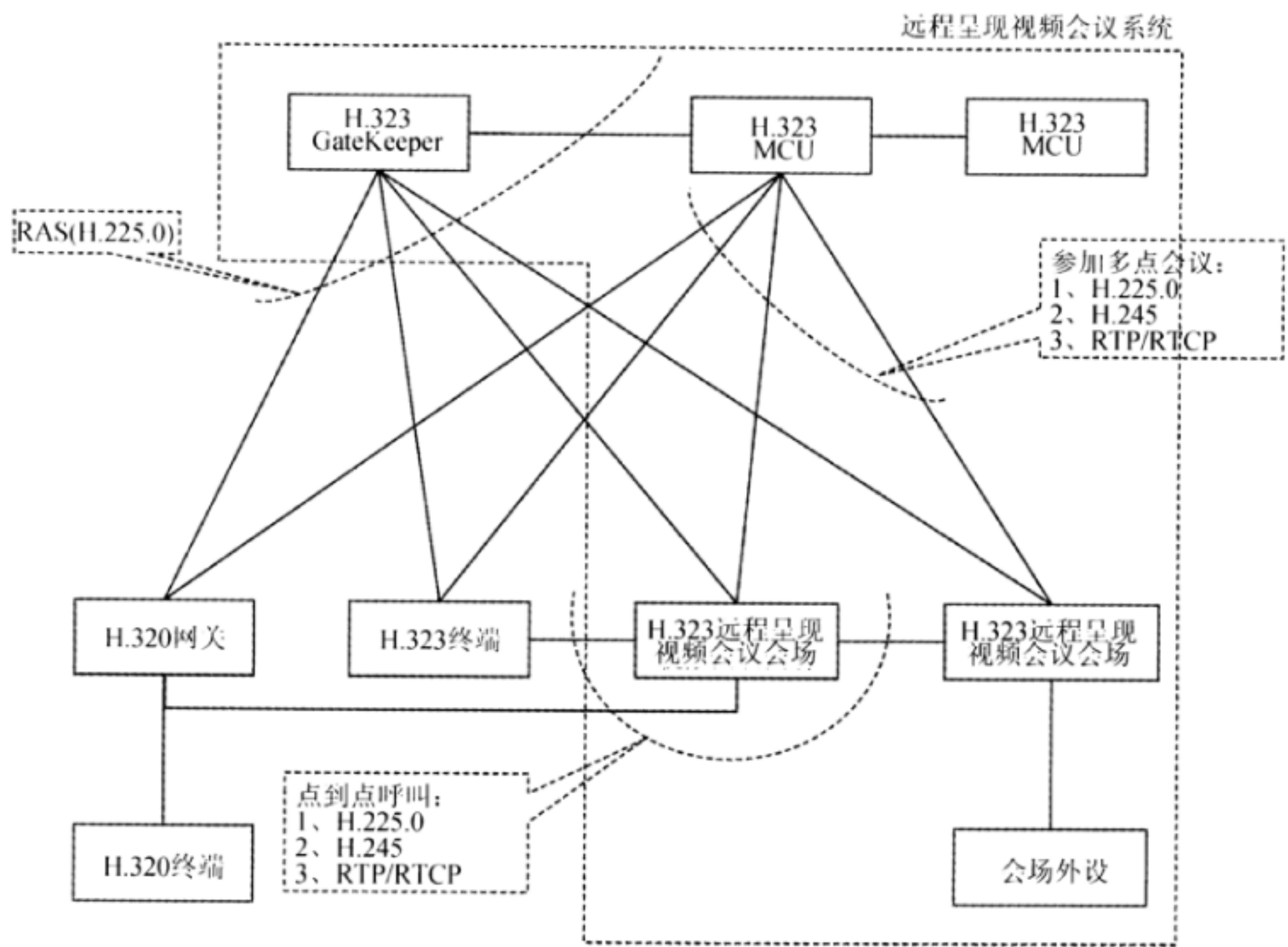


图 6 H.323 远程呈现信令模型

6.2.2 点对点呼叫流程

基于现有的 H.323 系统，点对点呼叫流程如图 7 所示。

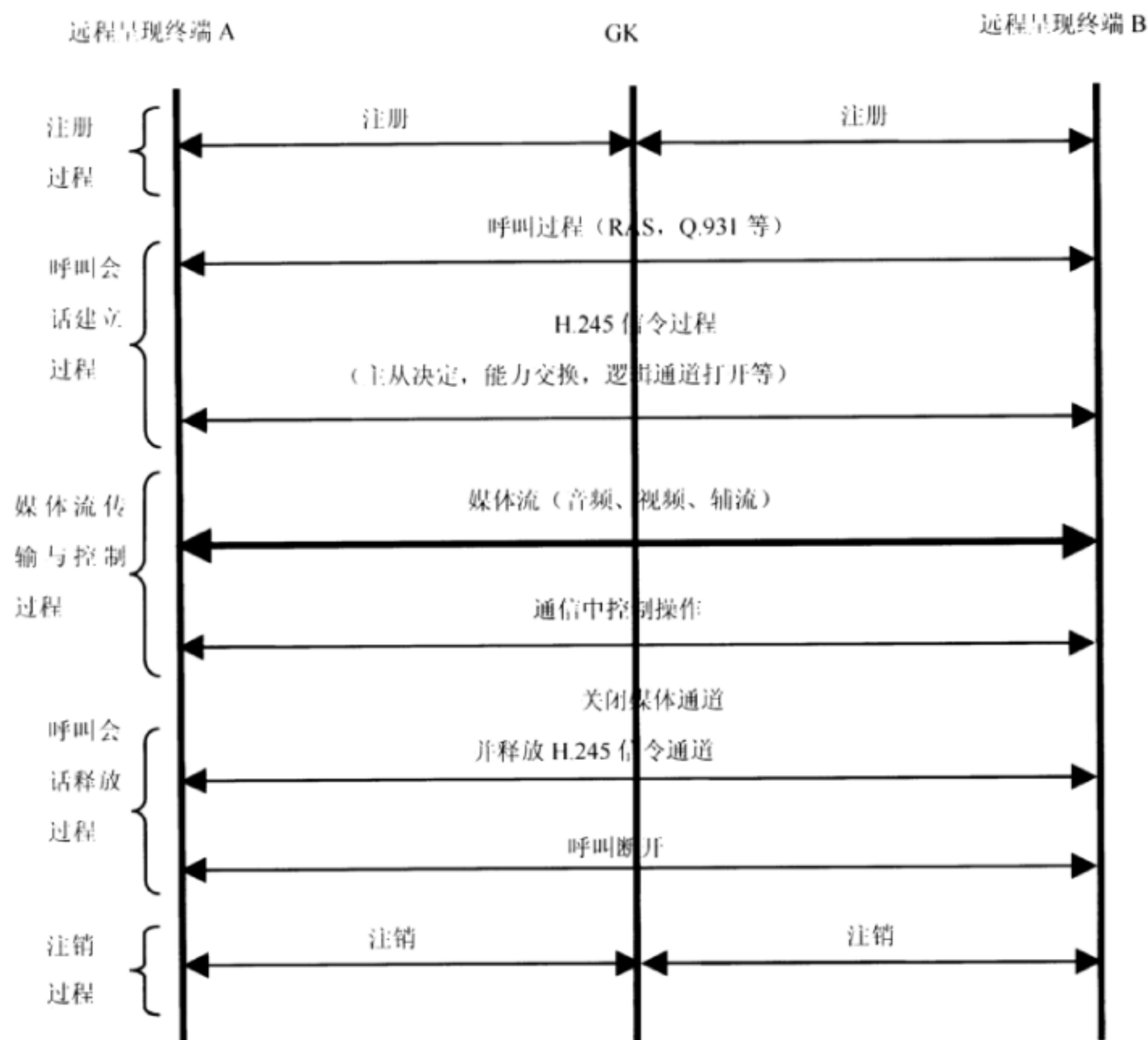


图 7 H.323 远程呈现点对点呼叫流程

整个呼叫流程包含下列步骤:

- a) 注册过程;
- b) 呼叫会话建立过程;
- c) 媒体流传输与控制过程;
- d) 呼叫会话释放过程;
- e) 注销过程。

其中, 影响互操作性的主要步骤为步骤 b) 和步骤 c)。

步骤 b) 和步骤 c) 的详细过程如图 8 所示。图 8 中所示例的是通过 GK 的点对点呼叫, 多点呼叫与其类似, 故不再赘述。

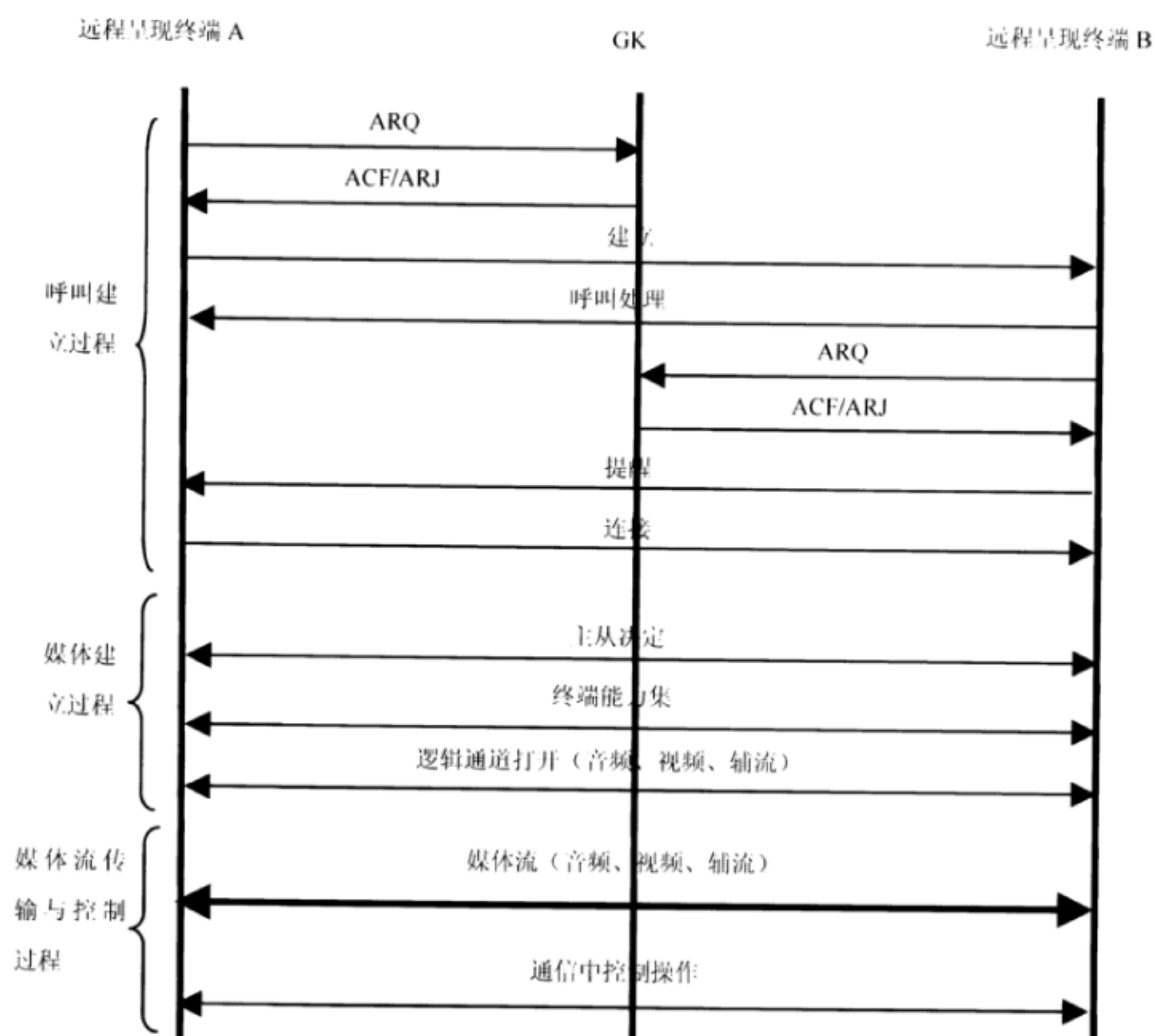


图 8 部分呼叫流程详解

如图 8 所示，详细的信令过程包括：

- a) 呼叫建立过程，包括 RAS 和 Q.931 呼叫信令。
- b) 媒体建立过程，主要使用 H.245 控制信令，包括主从决定、能力协商、逻辑通道打开。
- c) 媒体流传输与控制过程，包括媒体流传输与控制，使用 RTP/RTCP 和 H.245 控制信令。

6.2.3 多点呼叫流程

基于现有的 H.323 系统，多点呼叫流程与点对点流程类似，如图 9 所示。

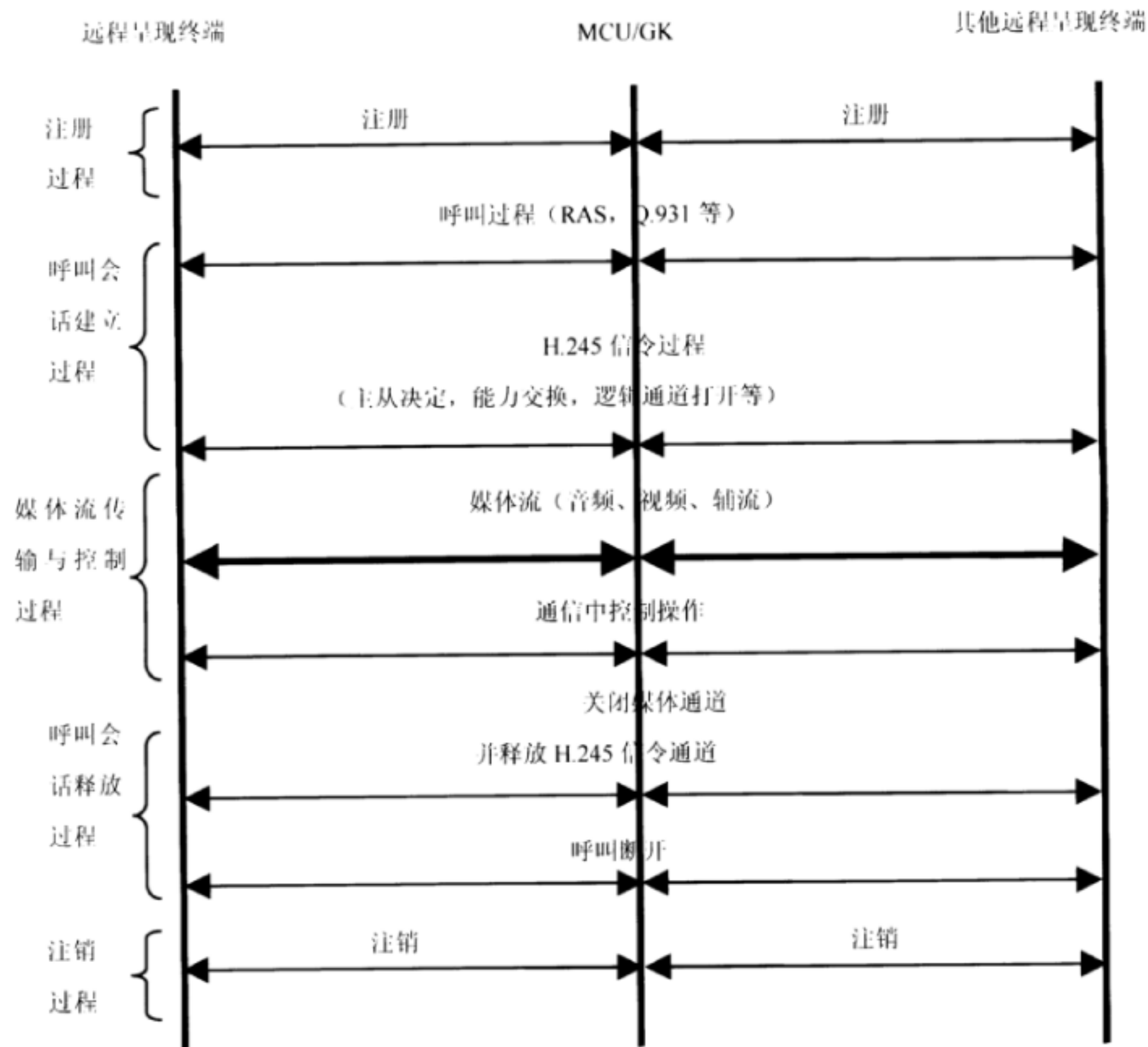


图 9 H.323 远程呈现多点呼叫流程

整个呼叫流程包含下列步骤：

- a) 注册过程；
- b) 呼叫会话建立过程；
- c) 媒体流传输与控制过程；
- d) 呼叫会话释放过程；
- e) 注销过程。

其中，影响互操作性的主要步骤为步骤 b) 和步骤 c)。其细节与点对点流程类似，故不再赘述。

6.3 SIP 体系

6.3.1 SIP 远程呈现信令模型

对 SIP 远程呈现视频会议系统，其涉及的协议/控制信令的内容如图 10 所示。

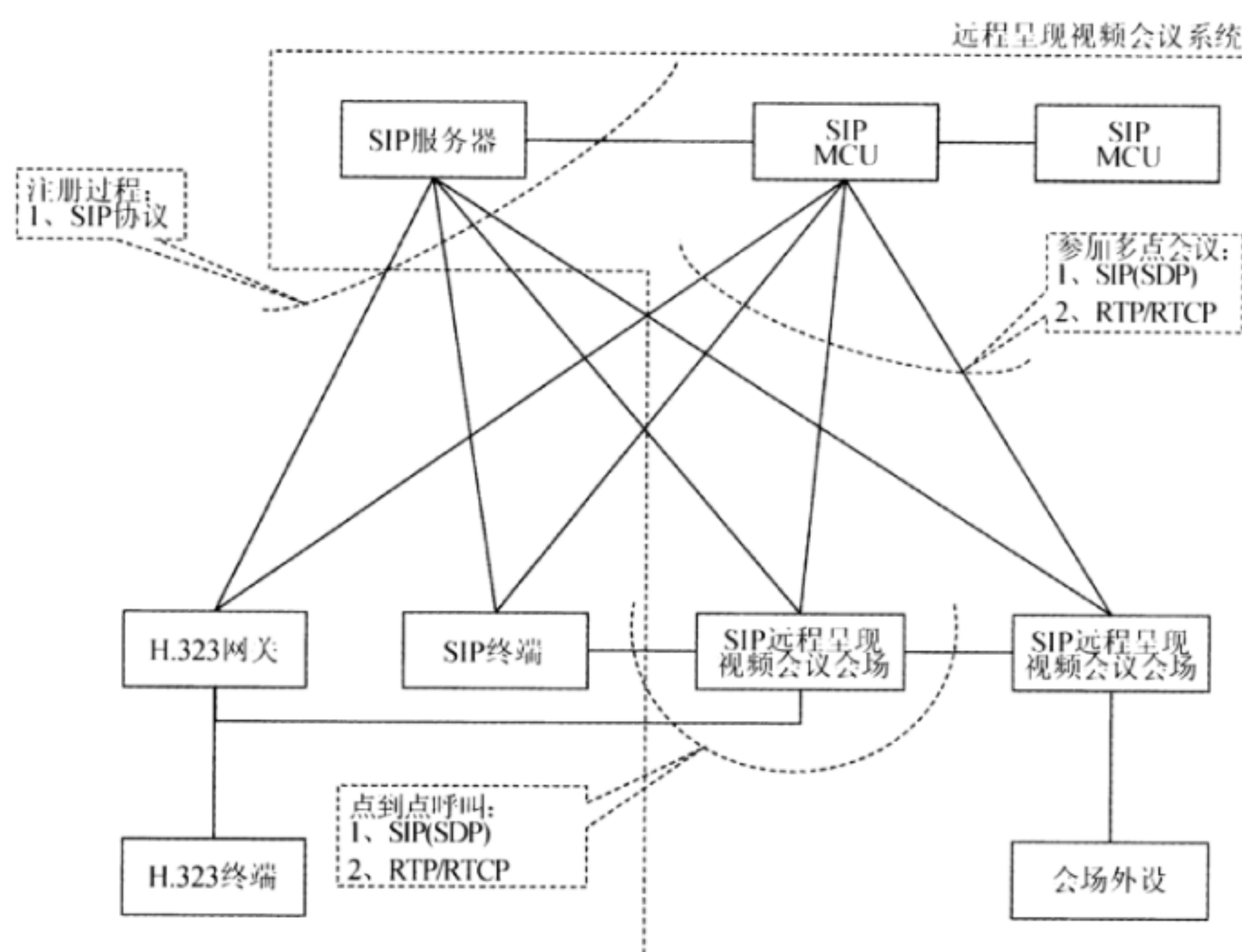


图 10 SIP 远程呈现信令模型

6.3.2 点对点呼叫流程

基于现有的 SIP 系统，点对点呼叫流程如图 11 所示。

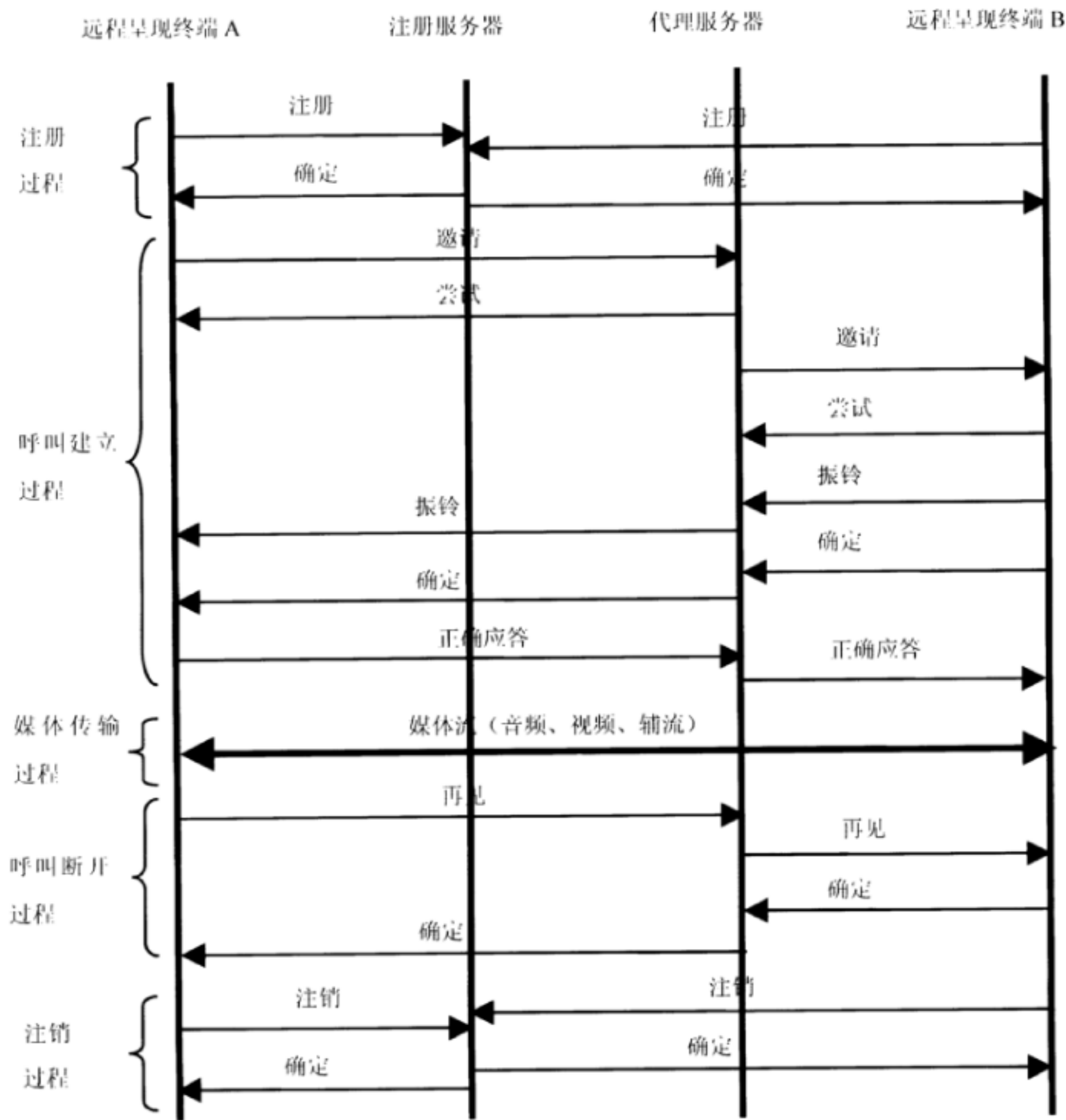


图 11 SIP 远程呈现呼叫流程

整个呼叫流程包含下列步骤：

- a) 注册过程（与 H.323 系统的注册过程对应）；
- b) 呼叫建立过程（与 H.323 系统的呼叫会话建立过程对应）；
- c) 媒体流传输过程（与 H.323 系统的媒体流传输与控制过程对应）；
- d) 呼叫断开过程（与 H.323 系统的呼叫会话释过程对应）；
- e) 注销过程（与 H.323 系统的注销过程对应）。

其中，影响互操作性的主要步骤为步骤 b）。

6.3.3 多点呼叫流程

基于现有的 SIP 系统，多点呼叫流程与点对点流程类似，如图 12 所示。

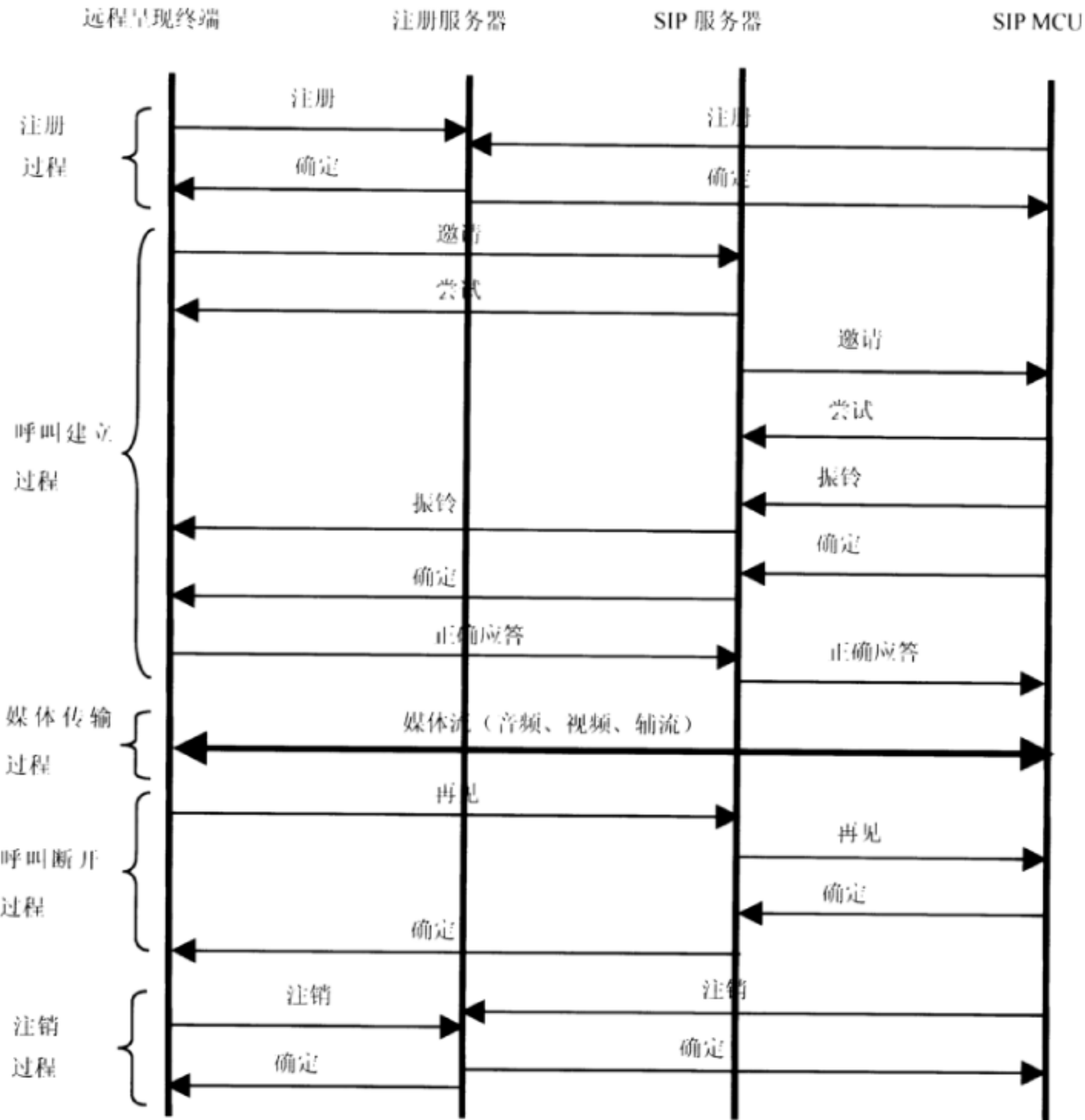


图 12 H.323 远程呈现多点呼叫流程

整个呼叫流程包含下列步骤：

- a) 注册过程（与 H.323 系统的注册过程对应）；
- b) 呼叫建立过程（与 H.323 系统的呼叫会话建立过程对应）；
- c) 媒体流传输过程（与 H.323 系统的媒体流传输与控制过程对应）；
- d) 呼叫断开过程（与 H.323 系统的呼叫会话释放过程对应）；
- e) 注销过程（与 H.323 系统的注销过程对应）。

其中，影响互操作性的主要步骤为步骤 b)。

6.4 H.323 与 SIP 互通

由于存在 H.323 和 SIP 协议，这必然会涉及 H.323 和 SIP 协议之间的互通考虑，以及分别基于这两种协议的远程呈现系统之间的互通，如：基于 H.323 协议的远程呈现终端和基于 SIP 协议的远程呈现终端之间的互通。如图 13 所示。

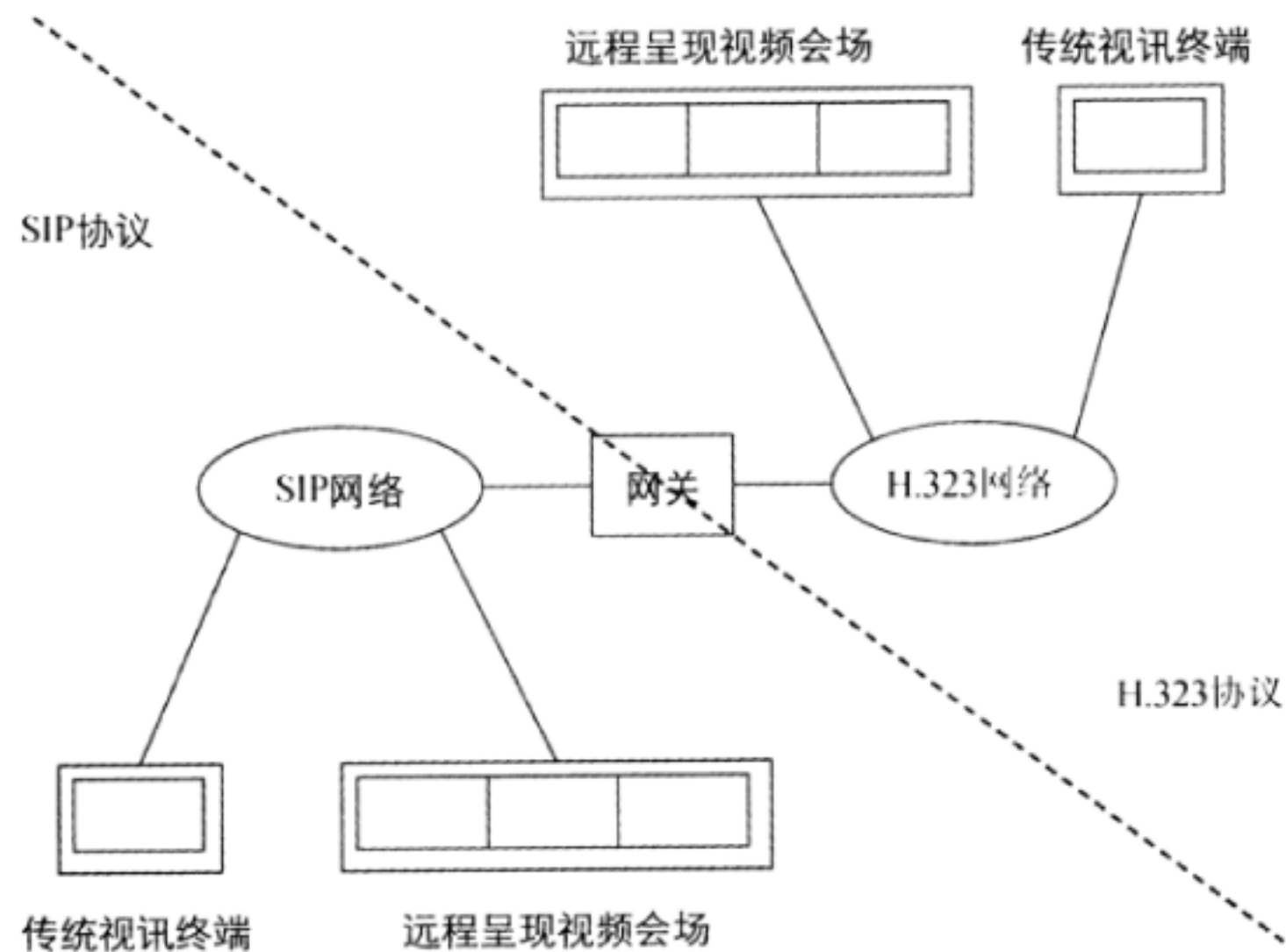


图 13 基于 H.323 的和基于 SIP 的远程呈现终端之间的互通

一个典型的例子，一个远程呈现视频会议终端位于 SIP 网络，而另一个远程呈现视频会议终端位于 H.323 网络。位于不同网络的远程呈现视频会议终端之间的典型呼叫流程如图 14 所示。

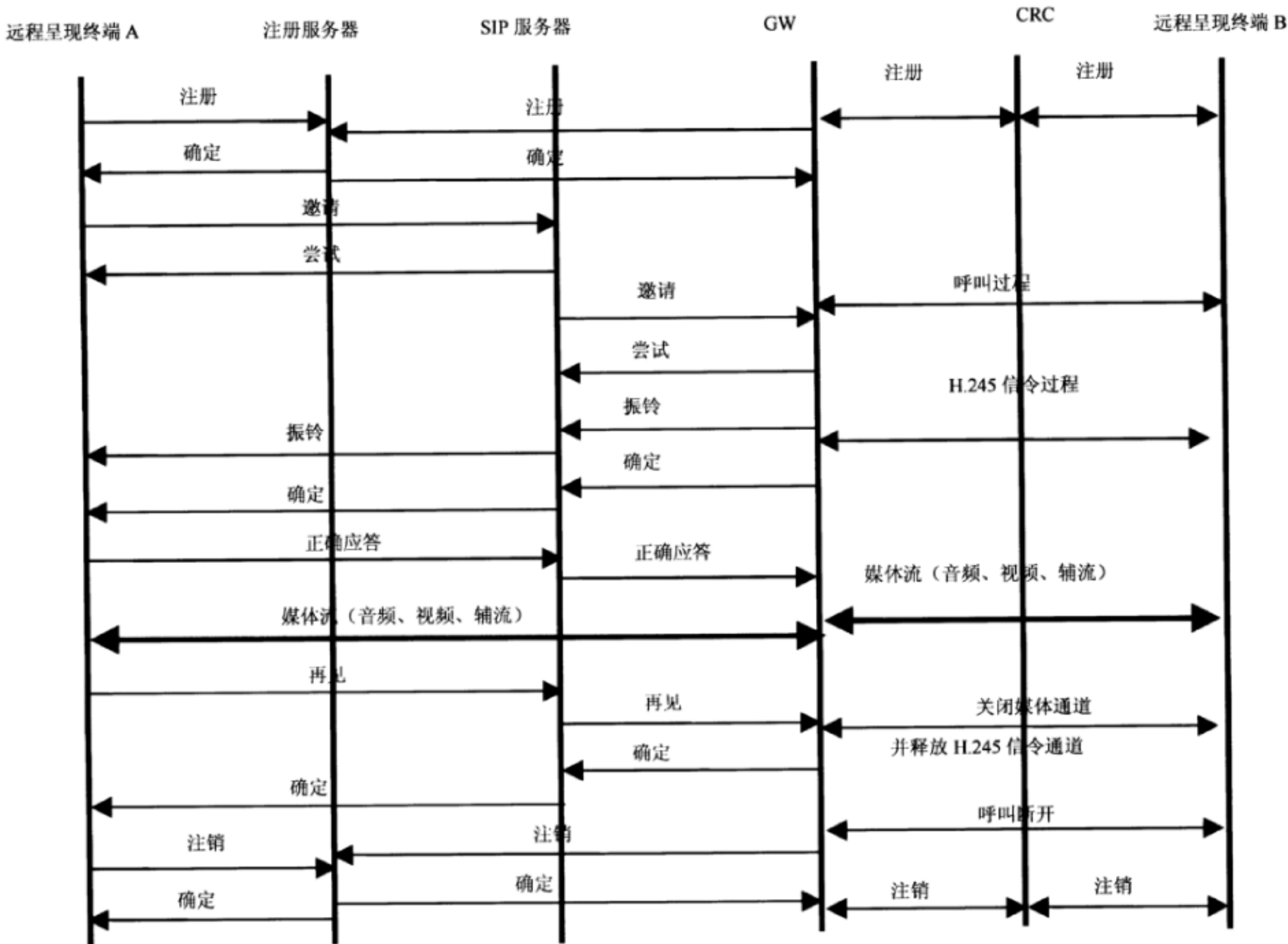


图 14 典型的例子，H.323 与 SIP 互通