

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

(43) 国际公布日  
2014年10月9日 (09.10.2014)



(10) 国际公布号  
WO 2014/161402 A2

- (51) 国际专利分类号: 无分类
- (21) 国际申请号: PCT/CN2014/072520
- (22) 国际申请日: 2014年2月25日 (25.02.2014)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201310673952.5 2013年12月11日 (11.12.2013) CN
- (71) 申请人: 中兴通讯股份有限公司 (ZTE CORPORATION) [CN/CN]; 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。
- (72) 发明人: 孙定准 (SUN, Dingzhun); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦中兴通讯股份有限公司转交, Guangdong 518057 (CN)。 冯斌 (FENG, Bin); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦中兴通讯股份有限公司转交, Guangdong 518057 (CN)。 朱存望 (ZHU, Cunwang); 中国广东省深圳市南山区

高新技术产业园科技南路中兴通讯大厦中兴通讯股份有限公司转交, Guangdong 518057 (CN)。

- (74) 代理人: 北京安信方达知识产权代理有限公司 (AFD CHINA INTELLECTUAL PROPERTY LAW OFFICE); 中国北京市海淀区学清路8号B座1601A, Beijing 100192 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH,

[见续页]

(54) Title: DISTRIBUTED VIDEO CONFERENCE METHOD, SYSTEM, TERMINAL, AND AUDIO-VIDEO INTEGRATED DEVICE

(54) 发明名称: 分布式视频会议实现方法、系统、终端、音视频一体化设备

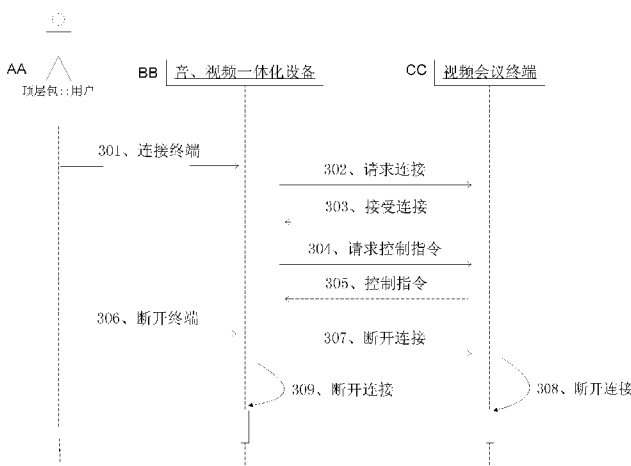


图3 / FIG. 3

- 301 CONNECTING TERMINAL
- 302 REQUESTING CONNECTION
- 303 ACCEPTING CONNECTION
- 304 REQUESTING CONTROL INSTRUCTION
- 305 CONTROL INSTRUCTION
- 306 DISCONNECTING TERMINAL
- 307 DISCONNECTING CONNECTION
- 308 DISCONNECTING CONNECTION
- 309 DISCONNECTING CONNECTION
- AA TOP LAYER PACKET::USER
- BB AUDIO-VIDEO INTEGRATED DEVICE
- CC VIDEO CONFERENCE TERMINAL

(57) Abstract: A distributed video conference method and system, terminal, and audio-video integrated device, relating to video conference technology. The audio-video integrated device comprises: a video collecting unit for collecting the video information at a conference site, including collecting the video information of the audio source location at the conference site; an audio collecting unit for collecting the audio information of the video conference; a network communication unit for transmitting the collected audio and video information to a video conference terminal. A video conference terminal and a distributed video conference method.

(57) 摘要: 一种分布式视频会议实现方法及系统、终端、音视频一体化设备, 涉及视频会议技术。所述音视频一体化设备, 包括: 视频采集单元, 采集会议现场的视频信息, 其中, 包括采集会议现场的声源位置的视频信息; 音频采集单元, 采集视频会议的音频信息; 网络通信单元, 将采集的音、视频信息传输给视频会议终端。一种视频会议终端和分布式视频会议的实现方法。



WO 2014/161402 A2



CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

**本国际公布:**

- 根据申请人的请求, 在条约第 21 条(2)(a)所规定的期限届满之前进行。
- 不包括国际检索报告, 在收到该报告后将重新公布(细则 48.2(g))。

**根据细则 4.17 的声明:**

- 关于申请人有权申请并被授予专利(细则 4.17(ii))
- 发明人资格(细则 4.17(iv))

## 分布式视频会议实现方法、系统、终端、音视频一体化设备

### 技术领域

5 本发明涉及视频会议，尤其涉及分布式视频会议实现方法及系统、终端、音视频一体化设备。

### 背景技术

10 随着视频摄像技术、网络宽带技术以及视频压缩技术的飞速发展，视频会议被广泛地使用于多种场合下的本地或远程的会议中。目前市场上，所见视频会议系统一般由视频会议终端，视频会议外部设备以及用于连接终端和外部设备的线缆组成，如图 1 所示。其中，视频会议终端负责协调整个会议系统各个单元的运行，包括音视频信号的编解码输出，音视频码流的打包解包以及与其他视频会议终端进行多媒体信息的交互等功能。视频会议外部设备一般包括音频、视频以及数据等多媒体信号采集、生成和显示设备，比如  
15 摄像头、麦克风、扬声器、电脑、电视、投影仪等。每种类型的外部设备一般可以是多个，而根据不同外部设备和会议终端之间的信息交互的关系，其布线方式多种多样。这样容易引起如下问题：

系统各个部分受到线缆长度的限制，无法移动相关外部设备，限制了设备的使用范围，从而约束了用户的活动范围。

20 设备布线的复杂，系统安装的调试易出错。

需要临时增加更多的外部设备，会增加布线的复杂程度。

由于将大部分的功能都集中在终端上，会增加终端的复杂度。

另外，对于原先的会议过程中，需要不断地根据不同场景和人群手动调整摄像头角度，造成一定的停滞，从而会影响会议的整体效果。

25

### 发明内容

本发明实施例提供一种分布式视频会议实现方法及系统、终端、音视频一体化设备，减少视频会议系统布线及终端的复杂度。

本发明实施例公开了一种音视频一体化设备，包括：

视频采集单元，其设置成采集会议现场的视频信息，其中，包括采集会议现场的声源位置的视频信息；

音频采集单元，其设置成采集视频会议的音频信息；以及

5 网络通信单元，其设置成将采集的音、视频信息传输给视频会议终端。

可选地，上述设备还包括：

视频编码单元，其设置成对采集的视频信息进行编码压缩；以及

音频编码单元，其设置成对采集的音频信息进行编码压缩。

10 可选地，上述设备中，所述网络通信单元，通过如下方式将采集的音、视频信息传输给视频会议终端：

所述网络通信单元，将编码压缩后的音、视频码流打包成媒体流实时传输协议格式并传输给视频会议终端。

可选地，上述设备还包括：

15 设备控制单元，其设置成对所述音频采集单元采集的音频信息进行分析，定位声源位置，控制所述视频采集单元采集所述声源位置的视频信息。

可选地，上述设备中，所述视频采集单元采用一个或一组摄像头。

可选地，上述设备中，所述视频采集单元采用一组摄像头时，所述设备控制单元根据所述声源位置来控制最佳位置的摄像头采集声源位置的图像，其余的摄像头分别采集会议现场不同区域的图像。

20 可选地，上述设备中，所述网络通信单元是设置成通过有线连接或者无线连接将所采集的音、视频信号传输给视频会议终端。

本发明实施例还公开了一种视频会议终端，包括：

网络通信单元，其设置成接收音视频一体化设备发送的音、视频码流包；

25 音频解码输出单元，其设置成对音视频一体化设备发送的音频码流进行解码，并将解码后的码流输出到输出设备；以及

视频解码输出单元，其设置成对音视频一体化设备发送的视频码流进行解码，并将解码后的码流输出到输出设备。

可选地，上述视频会议终端还包括：

音频编码单元,其设置成对音视频一体化设备发送的音频码流进行编码,再将编码后的音频码流发送给所述音频解码输出单元;以及

视频编码单元,其设置成对音视频一体化设备发送的视频码流进行编码,再将编码后的视频码流发送给所述视频解码输出单元。

5 可选地,上述视频会议终端还包括:

设备接入单元,其设置成根据用户操作,向所述音视频一体化设备发送视频采集控制指令,以控制所述音视频一体化设备采集用户需求的会场图像。

本发明实施例还公开了一种分布式视频会议系统,包括上述音视频一体化设备,以及上述视频会议终端。

10 本发明实施例还公开了一种分布式视频会议实现方法,包括:

音视频一体化设备采集会议现场的音、视频信息,将所采集的音、视频信息传输给视频会议终端,其中,所述音视频一体化设备采集的视频信息包括会议现场的声源位置的音频信息。

15 可选地,上述方法中,所述音视频一体化设备将所采集的音、视频信息传输给视频会议终端,包括:

所述音视频一体化设备将所采集的音、视频信息直接传输给视频会议终端;或者

20 所述音视频一体化设备对所采集的音、视频信息分别进行编码压缩,将编码压缩后的音、视频码流打包成媒体流实时传输协议格式并传输给视频会议终端。

可选地,上述方法中,所述音视频一体化设备采集的视频信息包括会议现场的声源位置的音频信息,包括:

所述音视频一体化设备采集音频信息时,还对所采集的音频信息进行分析,定位会议现场的声源位置,对所定位的声源位置进行视频采集。

25 可选地,上述方法中,所述音视频一体化设备采用一个或一组摄像头采集视频信息。

可选地,上述方法中,所述音视频一体化设备采用一组摄像头时,根据所述声源位置来控制最佳位置的摄像头采集声源位置的图像,其余的摄像头分别采集会议现场不同区域的图像。

可选地，上述方法中，所述音视频一体化设备通过有线连接或者无线连接，将所采集的音、视频信号传输给视频会议终端。

可选地，上述方法还包括：

所述视频会议终端接收所述音视频一体化设备发送的音、视频码流包；

5 以及

对所述音、视频码流包中的音、视频码流分别进行解码后输出到输出设备。

可选地，上述方法还包括：

10 所述视频会议终端对所述音、视频码流包中的音、视频码流先分别进行编码，再对编码后的音、视频码流分别进行解码。

可选地，上述方法还包括：

所述视频会议终端，根据用户操作，向所述音视频一体化设备发送视频采集控制指令；以及

15 所述音视频一体化设备根据所述视频采集控制指令采集用户需求的会场图像。

20 本发明实施例使用的摄像头与麦克风阵列一体化设备能够减少布线数量，降低布线复杂度，便于对会议室的布局。将编码分布到不同的一体化设备上，视频会议终端作为解码和会议控制设备，提高了编解码效率。同时，能够使音视频的采集更接近用户，根据声源快速切换画面，提供多种会场场景效果，提高视频会议整体效果。

## 附图概述

图 1 为现有视频会议系统基本组成的结构示意图；

25 图 2 为本发明实施例提供的分布式视频会议系统结构示意图；

图 3 为本发明实施例提供的音视频一体化设备与视频会议终端网络连接流程图；

图 4 为本发明实施例提供的音视频一体化设备的示意图；

图 5 为本发明实施例提供的音视频一体化设备音视频信号处理流程图；

图 6 为本发明实施例提供的视频会议终端的示意图；

图 7 为本发明实施例提供的视频会议终端音视频数据处理流程图；

图 8 为本发明实施例提供的一个实施例的会议系统结构图；

5 图 9 为本发明实施例提供的另一个实施例的会议系统结构图。

## 本发明的较佳实施方式

下文将结合附图对本发明实施例的技术方案作详细说明。需要说明的是，在不冲突的情况下，本申请的实施例和实施例中的特征可以任意相互组合。

10

### 实施例 1

本实施例提供一种分布式视频会议系统，如图 2 所示，至少包括音、视频一体化设备，视频会议终端，还可以包括电脑和输出设备。

15 其中，音、视频一体化设备，用于音视频信号采集和与所述的视频会议终端交互等。

视频会议终端，用于音视频数据解码，音视频数据输出和与其他视频会议终端进行交互等。

电脑，用于控制视频会议外部设备，控制视频会议终端，传输其他音视频数据等。

20 输出设备，用于输出音视频数据。

上述分布式视频会议系统中的音视频一体化设备与视频会议终端网络连接过程如图 3 所示，包括以下步骤：

步骤 301，用户请求音视频一体化设备连接视频会议终端；

步骤 302，音视频一体化设备向视频会议终端发送连接请求；

25 步骤 303，视频会议终端接受连接请求；

步骤 304，音视频一体化设备向视频会议终端请求控制指令；

步骤 305，视频会议终端向音视频一体化设备发送控制指令；

- 步骤 306, 用户请求音视频一体化设备断开与视频会议终端连接;
- 步骤 307, 音视频一体化设备向视频会议终端发送断开连接请求;
- 步骤 308, 视频会议终端处理断开连接, 释放资源;
- 步骤 309, 音视频一体化设备处理断开连接, 释放资源。

5

下面具体介绍上述系统中的音视频一体化设备 40, 如图 4 所示, 该设备包括如下各单元:

视频采集单元 41, 采集会议现场的视频信息, 其中, 包括采集会议现场的声源位置的视频信息;

10 音频采集单元 42, 采集视频会议的音频信息;

网络通信单元 43, 将采集的音、视频信息传输给视频会议终端。

需要说明的是, 上述音视频一体化设备还可以具有编码功能, 此时, 还包括对采集的视频信息进行编码压缩的视频编码单元 44, 以及对采集的音频信息进行编码压缩的音频编码单元 45。

15 而网络通信单元与视频会议终端的网络通信单元进行有线或无线连接, 用于音频编码单元与视频会议终端的音频解码输出单元的连接, 以及用于视频编码单元与视频会议终端的视频解码输出单元连接。实际应用中, 在数据传输过程中, 是由上述网络通信单元, 将编码压缩后的音、视频码流打包成媒体流实时传输协议格式传输给视频会议终端即可。

20 另外, 音视频一体化设备还可以包括:

设备控制单元 46, 用于连接视频采集单元 41 与音频采集单元 42, 可以对采集到的音频信号进行分析, 定位声源位置, 再将控制信息传递至视频采集单元 41, 控制摄像头, 采集声源位置图像信息。

25 视频采集单元 41 则可以采用一个或一组摄像头。当视频采集单元 41 采用一组摄像头时, 设备控制单元 46 根据声源位置控制最佳位置的摄像头采集声源位置的图像, 即控制最佳位置的摄像头转动跟踪声音最大的发言人, 其余的摄像头分别采集会议现场不同区域的图像。

当然, 音视频一体化设备还可以包括会议控制单元 47, 该单元主要用于

会议的连接与挂断、视频会议设备的控制、会议的显示模式控制、通信协议选择、网络状况监控和系统版本升级等。

较佳地，上述音视频一体化设备处理音视频信号的过程如图 5 所示，包括以下步骤：

5        步骤 501，利用视频采集单元视频信号，利用音频采集单元采集音频信号；

          步骤 502，根据音频信号，定位声源，控制视频采集单元，采集声源的图像信号；

10        步骤 503，将处理后的音、视频信号分别编码为当前视频会议所需的音、视频编码格式；

          上述步骤 503 的操作是可选，即可以不编码而直接发送音、视频信号。

          步骤 504，将编码后的音、视频码流打包成媒体流实时传输协议格式；

          步骤 505，将打包好的音、视频码流通过网络通信单元发送至视频会议终端。

15

          下面再介绍上述系统中的视频会议终端 60，如图 6 所示，包括如下单元：

          网络通信单元 61，接收音视频一体化设备发送的音、视频码流包；

          音频解码输出单元 62，对音视频一体化设备发送的音频码流进行解码，并将解码后的码流输出到输出设备；

20        视频解码输出单元 63，对音视频一体化设备发送的视频码流进行解码，并将解码后的码流输出到输出设备。

          需要说明的是，上述视频会议终端还可能需要具备编码功能，此时还包括一音频编码单元 64，对音视频一体化设备发送的音频码流进行编码，再将编码后的音频码流发送给所述音频解码输出单元 62。以及一视频编码单元  
25        65，对音视频一体化设备发送的视频码流进行编码，再将编码后的视频码流发送给所述视频解码输出单元 63。

          还有一些优选方案，在上述视频会议终端的架构基础上，增加设备接入单元 66，该单元根据用户操作，向音视频一体化设备发送视频采集控制指令，

以控制音视频一体化设备采集用户需求的会场图像。较佳地，可控制音视频一体化设备的摄像头转动以跟踪采集会场声音最大的发言人的图像。

较佳地，上述视频会议终端处理音视频数据的过程如图 7 所示，包括以下步骤：

5           步骤 701，接受经过视频会议终端的网络通信单元接收由音视频一体化设备发送来的音、视频码流包；

          步骤 702，如果在线进行视频会议，将接收到的音、视频码流包经所述的视频会议终端的网络通信单元发送给其他视频会议终端；

10          步骤 703，将接收到的音、视频码流包进行解包处理，得到当前视频会议所需的音、视频编码格式码流；

          步骤 704，将解包后的音、视频编码格式码流经音、视频解码输出单元进行解码输出；

          步骤 705，将解码后的音、视频码流发送给输出设备进行输出显示。

## 15          实施例 2

本实施例提供一种分布式视频系统，如图 8 所示，包括音视频一体化设备和视频会议终端。

音视频一体化设备，包括视频采集单元、音频采集单元、控制单元单元、音频编码单元、视频编码单元和网络通信单元等。

20          其中，音视频一体化设备中各个单元的具体介绍可参见实施例 1 的相应内容，在此不再赘述。

          视频会议终端，包括设备接入单元、会议控制单元、音频解码输出单元、视频解码输出单元和网络通信单元等。

25          其中，视频会议终端中各个单元的具体介绍可参见实施例 1 的相应内容，在此不再赘述。

          本实施例是基于音视频一体化的设备，将视频的编码单元、音频的编码单元集成到一体化设备中，根据摄像头采集的视频信号进行编码和根据麦克风阵列采集的音频信号进行编码，再通过网络或者相关的线缆，把编码后的

数据传输到视频会议终端，由视频会议终端解码输出或转发其他终端。本实施能够较大程度地减少视频会议终端的复杂度，提高编码效率。

### 实施例 3

5 本实施例提供一种分布式视频系统，如图 9 所示，同样包括音视频一体化设备和视频会议终端。

音视频一体化设备，包括视频采集单元、音频采集单元、设备控制单元、网络通信单元等。

10 其中，音视频一体化设备中各个单元的具体介绍可参见实施例 1 的相应内容，在此不再赘述。

视频会议终端，包括设备接入单元、会议控制单元、音频编码单元，视频编码单元，音频解码输出单元、视频解码输出单元和网络通信单元等。

其中，视频会议终端中各个单元的具体介绍可参见实施例 1 的相应内容，在此不再赘述。

15 本实施例是保留传统视频会议终端的视频的编码单元和音频编码单元，对于本发明实施例提出的音视频一体化设备不再负责视频的编码处理和音频编码处理，只负责对音视频数据的采集，基于声源对发言人图像采集和通过无线网络发送音视频数据的功能。本实施例能够较大减少布线数量，降低布线复杂度和便于对原先系统布局的改造。

20

### 实施例 4

本实施例提供一种分布式视频会议的实现方法，包括：

25 音视频一体化设备采集会议现场的音、视频信息，将所采集的音、视频信息传输给视频会议终端，其中，音视频一体化设备采集的视频信息包括会议现场的声源位置的视频信息。

其中，音视频一体化设备将所采集的音、视频信息传输给视频会议终端时，可以将所采集的音、视频信息直接传输给视频会议终端，也可以对所采集的音、视频信息分别进行编码压缩压缩后，打包成媒体流实时传输协议格式传输给视频会议终端。

另外，音视频一体化设备采集音频信息时，还可以对所采集的音频信息进行分析，定位会议现场的声源位置，再对所定位的声源位置进行视频采集。

实际应用中，音视频一体化设备可以采用一个或一组摄像头来采集视频信息。当采用一组摄像头时，根据声源位置控制最佳位置的摄像头采集声源位置的图像，其余的摄像头分别采集会议现场不同区域的图像。

在上述方法的基础上，还包括视频会议终端的处理操作，具体如下：

视频会议终端接收音视频一体化设备发送的音、视频码流包；

对音、视频码流包中的音、视频码流分别进行解码后输出到输出设备。

需要说明的是，当音视频一体化设备将所采集的音、视频信息直接传输给视频会议终端时，视频会议终端还需要对音、视频码流包中的音、视频码流先分别进行编码，再对编码后的音、视频码流分别进行解码。

还有一些方案提出，视频会议终端还可以根据用户操作，向音视频一体化设备发送视频采集控制指令，这样，音视频一体化设备就可以根据视频采集控制指令采集用户需求的会场图像，以提高图像采集的可靠性，满足用户需求。

本领域普通技术人员可以理解上述方法中的全部或部分步骤可通过程序来指令相关硬件完成，所述程序可以存储于计算机可读存储介质中，如只读存储器、磁盘或光盘等。可选地，上述实施例的全部或部分步骤也可以使用一个或多个集成电路来实现。相应地，上述实施例中的各模块/单元可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。本申请不限制于任何特定形式的硬件和软件的结合。

以上所述，仅为本发明的较佳实例而已，并非用于限定本发明的保护范围。凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

## 工业实用性

本申请技术方案使用的摄像头与麦克风阵列一体化设备能够减少布线数

量，降低布线复杂度，便于对会议室的布局。将编码分布到不同的一体化设备上，视频会议终端作为解码和会议控制设备，提高了编解码效率。同时，能够使音视频的采集更接近用户，根据声源快速切换画面，提供多种会场场景效果，提高视频会议整体效果。

5

## 权 利 要 求 书

1、一种音视频一体化设备，包括：

视频采集单元，其设置成采集会议现场的视频信息，其中，包括采集会议现场的声源位置的视频信息；

5 音频采集单元，其设置成采集视频会议的音频信息；以及  
网络通信单元，其设置成将采集的音、视频信息传输给视频会议终端。

2、如权利要求 1 所述的设备，还包括：

视频编码单元，其设置成对采集的视频信息进行编码压缩；以及  
音频编码单元，其设置成对采集的音频信息进行编码压缩。

10 3、如权利要求 2 所述的设备，其中，所述网络通信单元通过如下方式将采集的音、视频信息传输给视频会议终端：

所述网络通信单元，将编码压缩后的音、视频码流打包成媒体流实时传输协议格式并传输给视频会议终端。

4、如权利要求 1 至 3 任一项所述的设备，还包括：

15 设备控制单元，其设置成对所述音频采集单元采集的音频信息进行分析，定位声源位置，控制所述视频采集单元采集所述声源位置的视频信息。

5、如权利要求 4 所述的设备，其中，所述视频采集单元采用一个或一组摄像头。

20 6、如权利要求 5 所述的设备，其中，所述视频采集单元采用一组摄像头时，所述设备控制单元根据所述声源位置来控制最佳位置的摄像头采集声源位置的图像，其余的摄像头分别采集会议现场不同区域的图像。

7、如权利要求 5 所述的设备，其中，所述网络通信单元是设置成通过有线连接或者无线连接将所采集的音、视频信号传输给视频会议终端。

8、一种视频会议终端，包括：

25 网络通信单元，其设置成接收音视频一体化设备发送的音、视频码流包；

音频解码输出单元，其设置成对音视频一体化设备发送的音频码流进行解码，并将解码后的码流输出到输出设备；以及

视频解码输出单元，其设置成对音视频一体化设备发送的视频码流进行解码，并将解码后的码流输出到输出设备。

5 9、如权利要求 8 所述的视频会议终端，还包括：

音频编码单元，其设置成对音视频一体化设备发送的音频码流进行编码，再将编码后的音频码流发送给所述音频解码输出单元；以及

视频编码单元，其设置成对音视频一体化设备发送的视频码流进行编码，再将编码后的视频码流发送给所述视频解码输出单元。

10 10、如权利要求 8 或 9 所述的视频会议终端，还包括：

设备接入单元，其设置成根据用户操作，向所述音视频一体化设备发送视频采集控制指令，以控制所述音视频一体化设备采集用户需求的会场图像。

11、一种分布式视频会议系统，其中，包括如权利要求 1 至 7 所述的音视频一体化设备，以及如权利要求 8 至 10 所述的视频会议终端。

15 12、一种分布式视频会议实现方法，包括：

音视频一体化设备采集会议现场的音、视频信息，将所采集的音、视频信息传输给视频会议终端，其中，所述音视频一体化设备采集的视频信息包括会议现场的声源位置的视频信息。

20 13、如权利要求 12 所述的方法，其中，所述音视频一体化设备将所采集的音、视频信息传输给视频会议终端，包括：

所述音视频一体化设备将所采集的音、视频信息直接传输给视频会议终端；或者

25 所述音视频一体化设备对所采集的音、视频信息分别进行编码压缩，将编码压缩后的音、视频码流打包成媒体流实时传输协议格式并传输给视频会议终端。

14、如权利要求 13 所述的方法，其中，所述音视频一体化设备采集的视

频信息包括会议现场的声源位置的视频信息，包括：

所述音视频一体化设备采集音频信息时，还对所采集的音频信息进行分析，定位会议现场的声源位置，对所定位的声源位置进行视频采集。

15、如权利要求 14 所述的方法，其中，

5 所述音视频一体化设备采用一个或一组摄像头采集视频信息。

16、如权利要求 15 所述的方法，其中，所述音视频一体化设备采用一组摄像头时，根据所述声源位置来控制最佳位置的摄像头采集声源位置的图像，其余的摄像头分别采集会议现场不同区域的图像。

17、如权利要求 13 至 16 任一项所述的方法，其中，

10 所述音视频一体化设备通过有线连接或者无线连接，将所采集的音、视频信号传输给视频会议终端。

18、如权利要求 13 至 16 任一项所述的方法，还包括：

所述视频会议终端接收所述音视频一体化设备发送的音、视频码流包；  
以及

15 对所述音、视频码流包中的音、视频码流分别进行解码后输出到输出设备。

19、如权利要求 18 所述的方法，还包括：

所述视频会议终端对所述音、视频码流包中的音、视频码流先分别进行编码，再对编码后的音、视频码流分别进行解码。

20 20、如权利要求 18 所述的方法，还包括：

所述视频会议终端，根据用户操作，向所述音视频一体化设备发送视频采集控制指令；以及

所述音视频一体化设备根据所述视频采集控制指令采集用户需求的会场图像。

25

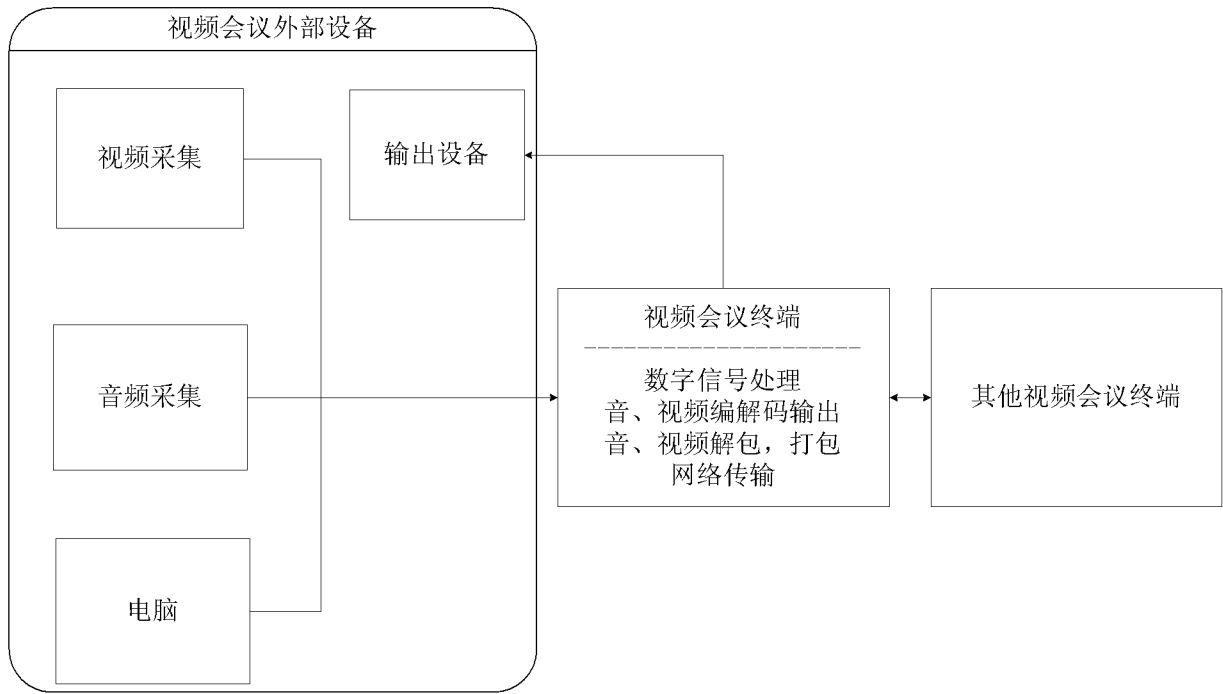


图 1

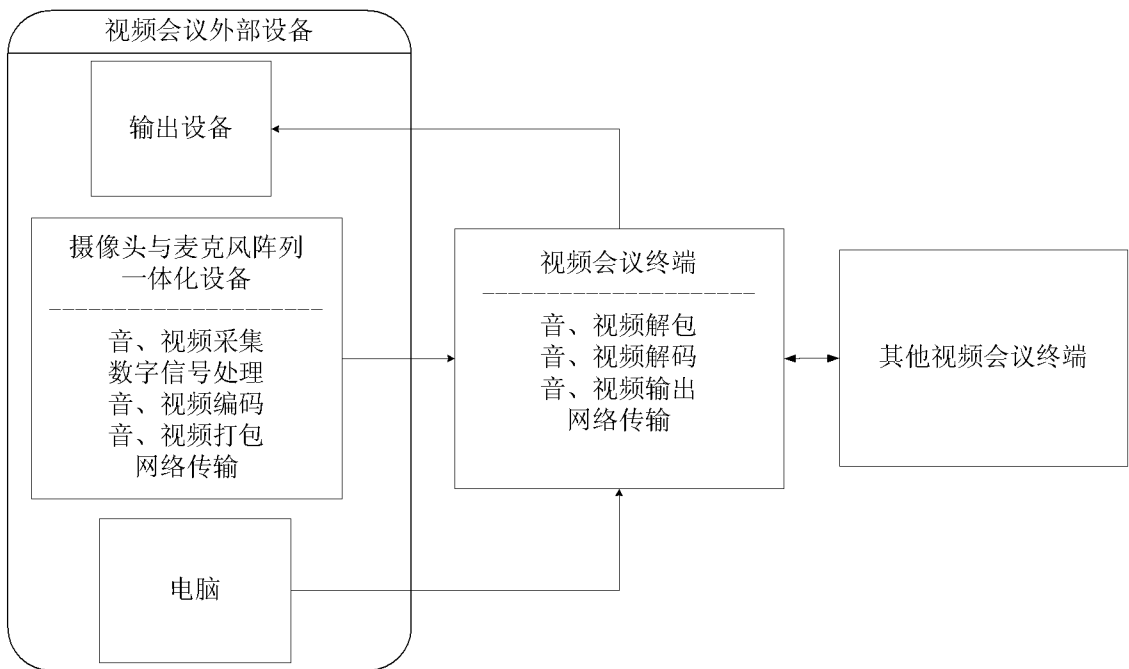


图 2

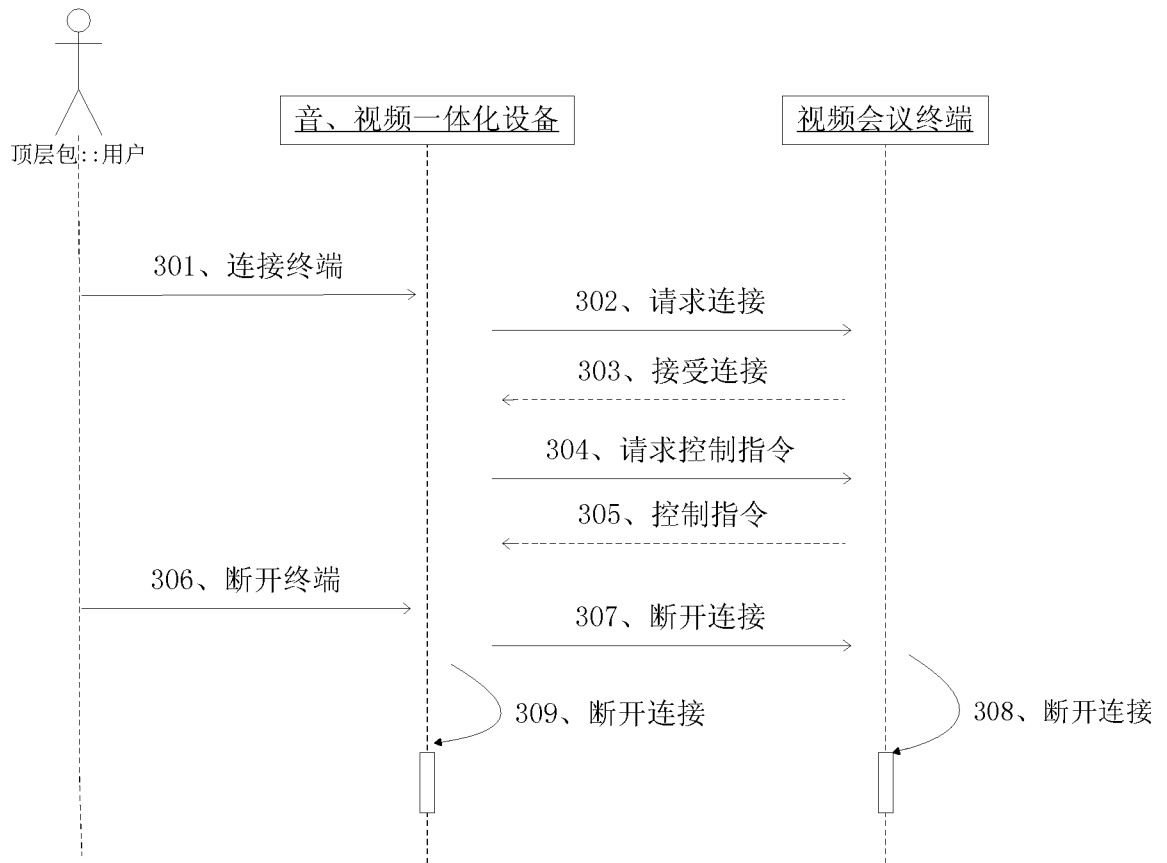


图 3



图 4

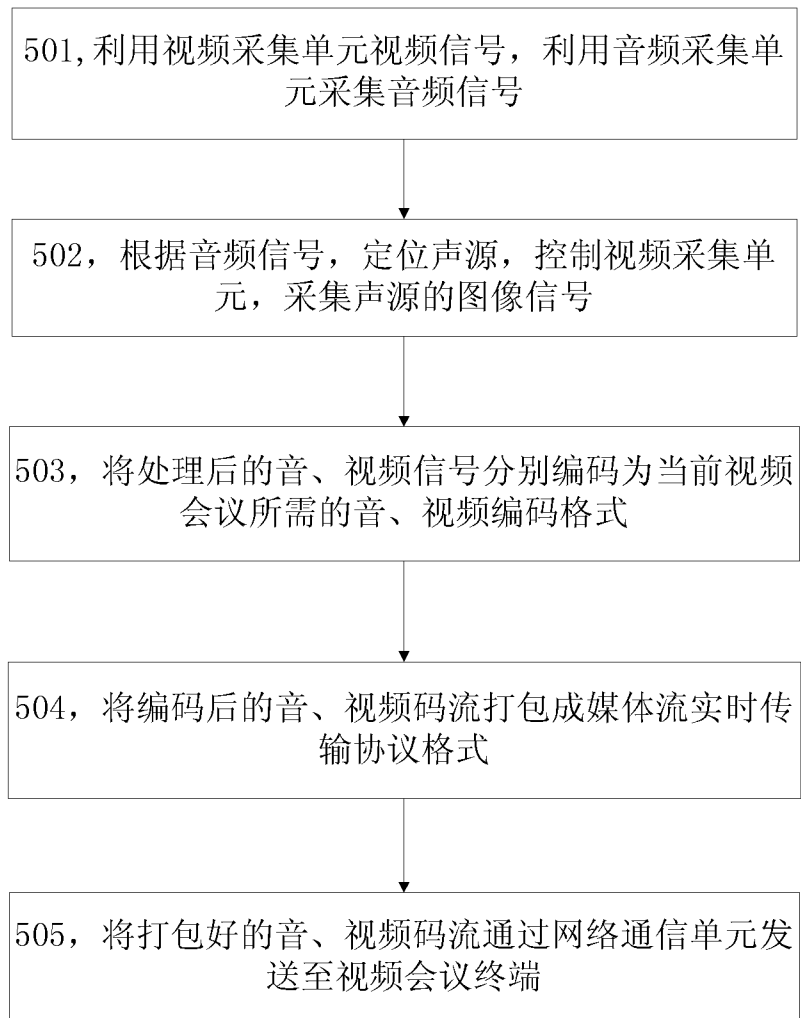


图 5



图 6

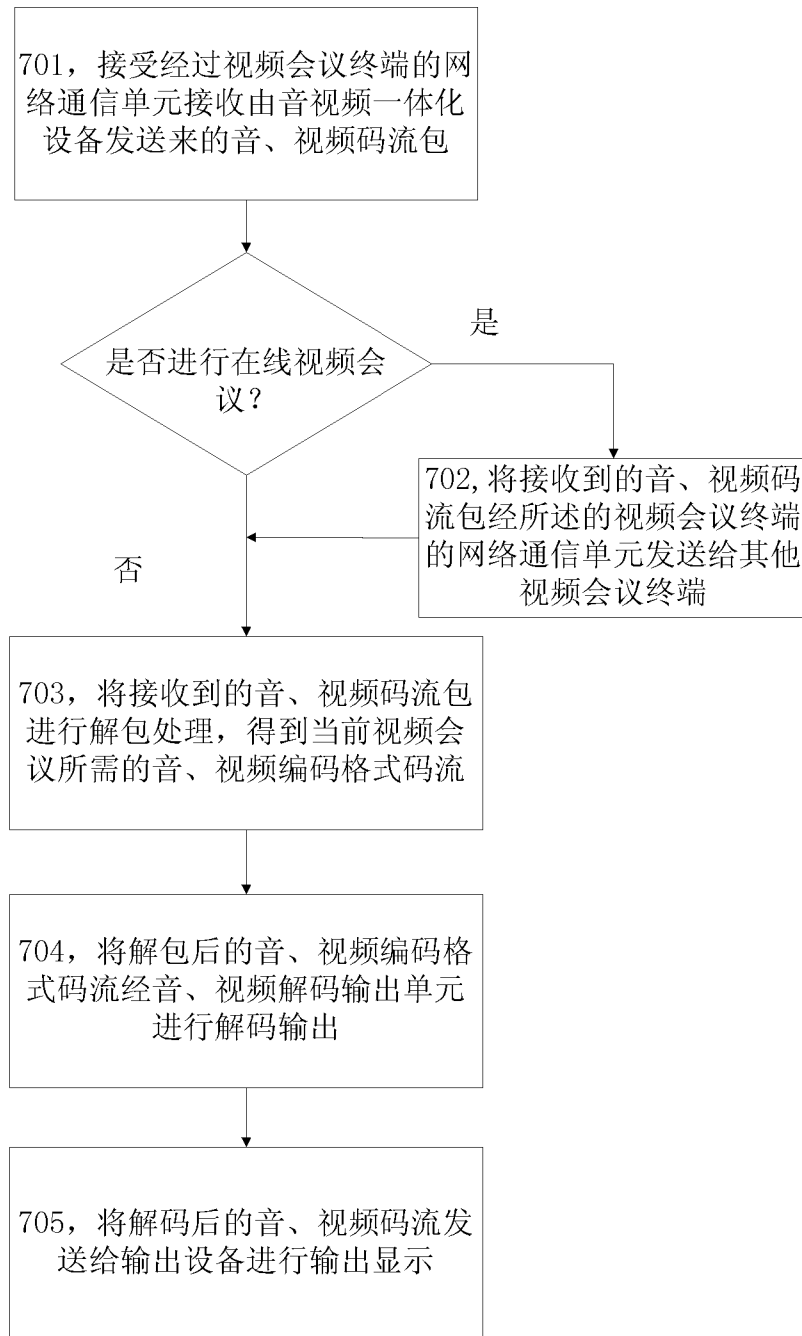


图 7

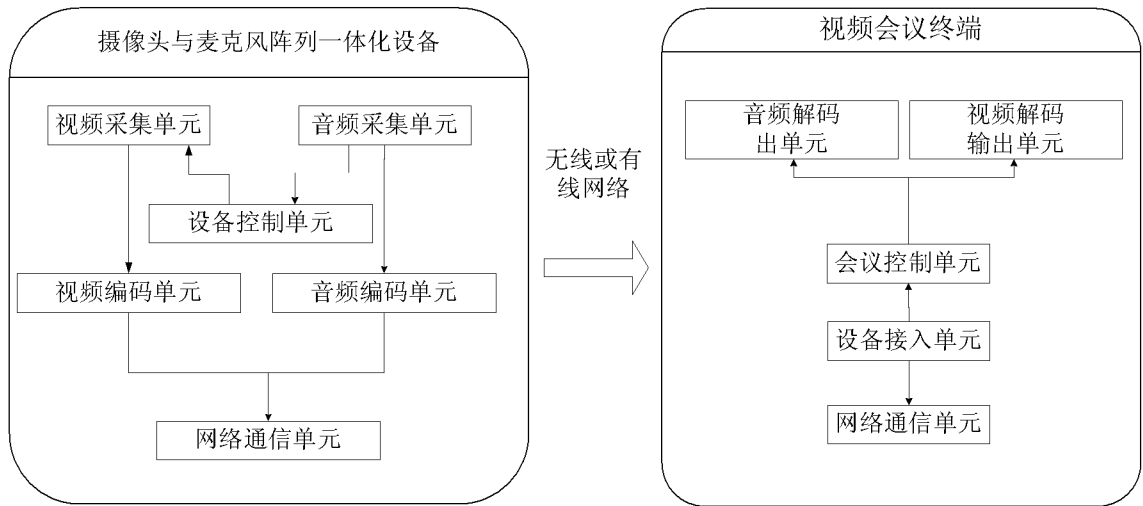


图 8

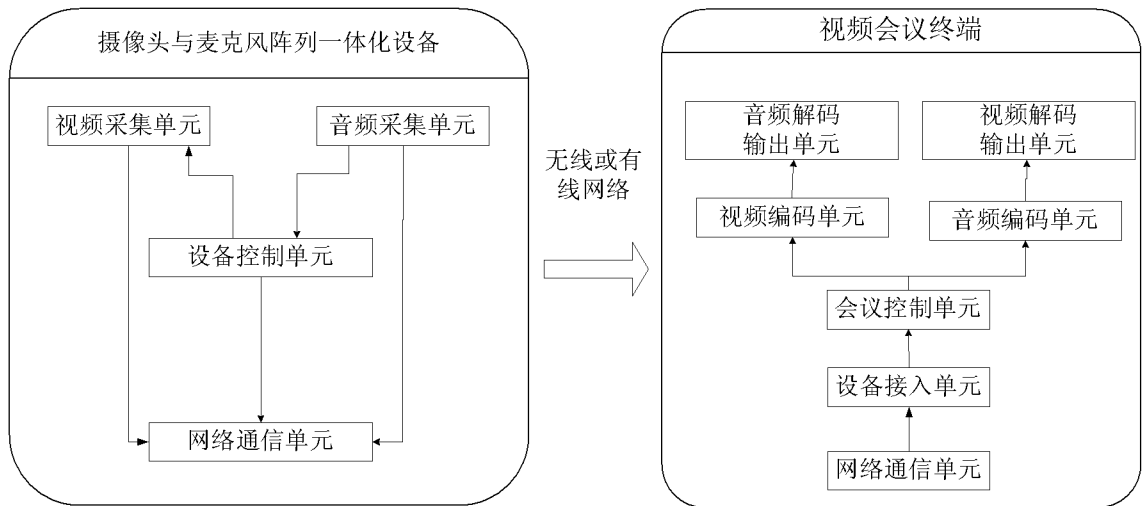


图 9