



(12) 发明专利申请

(10) 申请公布号 CN 101827271 A

(43) 申请公布日 2010.09.08

(21) 申请号 200910046978.0

(22) 申请日 2009.03.04

(71) 申请人 联芯科技有限公司

地址 200233 上海市钦江路 333 号 41 幢 4 楼

(72) 发明人 朱胜 裘风光

(74) 专利代理机构 上海专利商标事务所有限公司 31100

代理人 骆希聪

(51) Int. Cl.

H04N 7/52(2006.01)

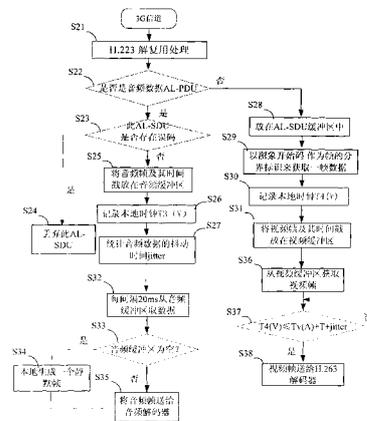
权利要求书 2 页 说明书 6 页 附图 4 页

(54) 发明名称

音频视频同步方法、装置以及数据接收终端

(57) 摘要

本发明涉及一种音频视频同步方法、其装置，以及实现音频视频同步的数据接收终端。该方法包括：于数据发送端，在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳，以及于数据接收终端，执行以下步骤：获取接收的音频数据及相应的音频时间戳，并记录本地当前时钟；获取接收的视频数据及相应的视频时间戳，形成完整的视频数据帧，并记录本地当前时钟；依据音频数据帧的音频时间戳及本地当前时钟计算音频抖动时间；每隔预定时间将一帧音频数据送给音频解码器，如果音频数据已被丢弃，则生成一静默帧送给音频解码器；参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间，以决定是否将此视频数据帧交给视频解码器。



1. 一种音频视频同步方法,是应用于低码率无线信道在高误码率环境下的音频视频同步,该方法包括:

于数据发送端,在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳;以及

于数据接收终端,执行步骤:

获取接收的音频数据及相应的音频时间戳,并记录本地当前时钟;

获取接收的视频数据及相应的视频时间戳,形成完整的视频数据帧,并记录本地当前时钟;

依据音频数据帧的音频时间戳及本地当前时钟计算音频抖动时间;

每隔预定时间将一帧音频数据送给音频解码器,如果音频数据因误码而被丢弃,则生成一静默帧送给音频解码器;以及

参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器。

2. 如权利要求 1 所述的方法,其特征在于,于数据发送端,在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳的步骤进一步包括:

记录本地当前视频时间戳;

将一帧视频数据分割成多个视频数据单元;

将各视频数据单元形成视频数据包,其中各视频数据包包含所述视频时间戳;

记录本地当前音频时间戳;

将一帧音频数据作为一音频数据单元,并将该音频数据单元形成音频数据包,其中该音频数据包包含所述音频时间戳;

对各视频数据包及该音频数据包进行复用处理,并发送给数据接收端。

3. 如权利要求 1 所述的方法,其特征在于,参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器的步骤包括:

如果该待处理的视频帧的视频时间戳大于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则不将视频帧交给视频解码器;

如果该待处理的视频帧的视频时间戳小于等于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则将视频帧交给视频解码器。

4. 一种音频视频同步装置,是应用于低码率无线信道在高误码率环境下的音频视频同步,该装置包括:

数据发送端,在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳;以及

数据接收终端,包括:

获取接收的音频数据及相应的音频时间戳,并记录本地当前时钟的单元;

获取接收的视频数据及相应的视频时间戳,形成完整的视频数据帧,并记录本地当前时钟的单元;

依据相邻音频数据帧的音频时间戳及本地当前时钟计算所述音频抖动时间的单元;

每隔预定时间将一帧音频数据送给音频解码器,如果音频数据因误码而被丢弃,则生

成一静默帧送给音频解码器的单元；以及

参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间，以决定是否将此视频数据帧交给视频解码器的单元。

5. 如权利要求 4 所述的装置，其特征在于，在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳进一步包括：

记录本地当前视频时间戳；

将一帧视频数据分割成多个视频数据单元；

将各视频数据单元形成视频数据包，其中各视频数据包包含所述视频时间戳；

记录本地当前音频时间戳；

将一帧音频数据作为一音频数据单元，并将该音频数据单元形成音频数据包，其中该音频数据包包含所述音频时间戳；

对各视频数据包及该音频数据包进行复用处理，并发送给数据接收端。

6. 如权利要求 4 所述的装置，其特征在于，在参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间，以决定是否将此视频数据帧交给视频解码器的单元中，如果该待处理的视频帧的视频时间戳大于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值，则不将视频帧交给视频解码器；如果该待处理的视频帧的视频时间戳小于等于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值，则将视频帧交给视频解码器。

7. 一种数据接收终端，用以接收数据发送端的视频和音频数据并同步地播放，该数据发送端在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳，其中该数据接收终端包括：

获取接收的音频数据及相应的音频时间戳，并记录本地当前时钟的单元；

获取接收的视频数据及相应的视频时间戳，形成完整的视频数据帧，并记录本地当前时钟的单元；

依据相邻音频数据帧的音频时间戳及本地当前时钟计算所述音频抖动时间的单元；

每隔预定时间将一帧音频数据送给音频解码器，如果音频数据因误码而被丢弃，则生成一静默帧送给音频解码器的单元；以及

参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间，以决定是否将此视频数据帧交给视频解码器的单元。

8. 如权利要求 7 所述的数据接收终端，其特征在于，在参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间，以决定是否将此视频数据帧交给视频解码器的单元中，如果该待处理的视频帧的视频时间戳大于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值，则不将视频帧交给视频解码器；如果该待处理的视频帧的视频时间戳小于等于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值，则将视频帧交给视频解码器。

## 音频视频同步方法、装置以及数据接收终端

### 技术领域

[0001] 本发明涉及低码率无线信道传输,尤其是涉及低码率无线信道终端在高误码率环境下的音频视频同步方法及其装置,以及用于实现视频音频同步的数据接收终端。

### 背景技术

[0002] 低码率无线信道终端在高误码率环境下,因为发送端、网络、接收端的处理而引起音频视频不同步,包含主要原因为:1、由于发送端对视频码率的控制会导致视频帧率引起相应的变化;2、由于网络抖动或网络误码会引起音频视频数据时序的变化;3、由于接收端对音频视频数据的缓存而延迟播放;这三个原因在实际开发过程中都是实际存在的,然而这些原因导致的音频视频不同步并没有得到有效地解决。

[0003] H. 324/M 国际标准可以在无线电路交换网络中支持实时多媒体服务应用。该标准包含的几个子协议标准是:语音、视频、用户数据和控制数据的多路复用和分离(H. 223);3GPP 采纳 H. 324/M 建议作为 3G 网络传统视频电话的一个标准,被其采纳的建议被命名为 3G-324M,3G-324M 终端是应用于无线电路交换网络的视频、音频和数据的实时传输设备,但是它针对语音、视频和多路复用操作提出了一些要求,如:它指定 H. 263 作为强制(视频编码)基本标准,而把 MPEG-4 作为视频编码推荐标准;指定 AMR 作为强制音频编码标准,而把 G. 732. 1 作为音频编码推荐标准;添加 H. 223 附件 B 用来保护复用数据。

[0004] 3G 业务中规定误码率可以达到  $10^{-4}$  到  $10^{-6}$ ,在信号质量比较差的情况下,误码率将达到  $10^{-3}$ ,所以因为误码原因,会使视频质量下降。同时目前 3G-324M 业务中没有使用音频视频同步措施,如果通话时间过长,用户会明显感觉出音频视频不同步。

[0005] 下面以 videophone 业务为例介绍现有技术存在的问题:

[0006] 根据 3G324 协议规定,建议采用 CS 64K 信道传输,视频数据的码率为 48kbps 左右、帧率在 5 ~ 15 帧/秒、音频数据的码率为 12kbps 左右、帧率为 50 帧/秒方案,目前是通过 H. 245 中失真指示 H223SkewIndication 来实现同步的,即(可以参考 ITU-T H. 245):

[0007] H223SkewIndication:: = SEQUENCE

[0008] {

[0009] logicalChannelNumber1 LogicalChannelNumber,

[0010] logicalChannelNumber2 LogicalChannelNumber,

[0011] skew INTEGER(0.. 4095), -- 单位 ms

[0012] }

[0013] 此失真指示用于指示远程终端视频逻辑信道和音频逻辑信道间时间失真的平均值。其中 logicalChannelNumber1 和 logicalChannelNumber2 均为处于打开状态的逻辑信道的信道编号。失真信息包括采样时间、编码器时延以及发送端缓存器时延的差异,并且失真以表示给定采样点数据的首比特传输时间为基准度量。此失真信息不包括网络抖动或网络误码会引起音频视频数据时序变化的信息。

[0014] H. 245 中的失真指示 H223SkewIndication 只包括发送端音频视频采样时间、编码

器编码时延以及发送端的缓存器时延的信息,它表明只是由于发送端的原因而引起的差异才导致音频视频不同步,并没有包括由于网络抖动或网络误码会引起音频视频数据时序的变化和接收端对音频视频数据的缓存信息,这种方法只能部分地解决音频视频同步现象,不能彻底地解决不同步问题。由于视频的帧率是在不断变化、存在网络抖动或网络误码、接收端存在数据缓冲等原因,进行一段时间的视频通话后,仍然会导致音频视频不同步。

## 发明内容

[0015] 本发明所要解决的技术问题是提供一种低码率无线信道在高误码率环境下的音频视频同步方法、其装置,以及实现音频视频同步的数据接收终端。

[0016] 本发明为解决上述技术问题而采用的技术方案是提出一种音频视频同步方法,是应用于低码率无线信道在高误码率环境下的音频视频同步,该方法包括:

[0017] 于数据发送端,在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳;以及

[0018] 于数据接收终端,执行步骤:

[0019] 获取接收的音频数据及相应的音频时间戳,并记录本地当前时钟;

[0020] 获取接收的视频数据及相应的视频时间戳,形成完整的视频数据帧,并记录本地当前时钟;

[0021] 依据音频数据帧的音频时间戳及本地当前时钟计算所述音频抖动时间;

[0022] 每隔预定时间将一帧音频数据送给音频解码器,如果音频数据因误码而被丢弃,则生成一静默帧送给音频解码器;以及

[0023] 参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器。

[0024] 在本发明的一实施例中,于数据发送端,在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳的步骤进一步包括:

[0025] 记录本地当前视频时间戳;

[0026] 将一帧视频数据分割成多个视频数据单元;

[0027] 将各视频数据单元形成视频数据包,其中各视频数据包包含所述视频时间戳;

[0028] 记录本地当前音频时间戳;

[0029] 将一帧音频数据作为一音频数据单元,并将该音频数据单元形成音频数据包,其中该音频数据包包含所述音频时间戳;

[0030] 对各视频数据包及该音频数据包进行复用处理,并发送给数据接收端。

[0031] 在本发明的一实施例中,参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器的步骤包括:

[0032] 如果该待处理的视频帧的视频时间戳大于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则不将视频帧交给视频解码器;

[0033] 如果该待处理的视频帧的视频时间戳小于等于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则将视频帧交给视频解码器。

[0034] 本发明另提供一种用以执行上述方法的音频视频同步装置,该装置包括:

[0035] 数据发送端,在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳;以及

[0036] 数据接收终端,包括:

[0037] 获取接收的音频数据及相应的音频时间戳,并记录本地当前时钟的单元;

[0038] 获取接收的视频数据及相应的视频时间戳,形成完整的视频数据帧,并记录本地当前时钟的单元;

[0039] 依据相邻音频数据帧的音频时间戳及本地当前时钟计算所述音频抖动时间的单元;

[0040] 每隔预定时间将一帧音频数据送给音频解码器,如果音频数据因误码而被丢弃,则生成一静默帧送给音频解码器的单元;以及

[0041] 参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器的单元。

[0042] 在本发明的一实施例中,在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳进一步包括:

[0043] 记录本地当前视频时间戳;

[0044] 将一帧视频数据分割成多个视频数据单元;

[0045] 将各视频数据单元形成视频数据包,其中各视频数据包包含所述视频时间戳;

[0046] 记录本地当前音频时间戳;

[0047] 将一帧音频数据作为一音频数据单元,并将该音频数据单元形成音频数据包,其中该音频数据包包含所述音频时间戳;

[0048] 对各视频数据包及该音频数据包进行复用处理,并发送给数据接收端。

[0049] 在本发明的一实施例中,在参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器的单元中,如果该待处理的视频帧的视频时间戳大于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则不将视频帧交给视频解码器;如果该待处理的视频帧的视频时间戳小于等于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则将视频帧交给视频解码器。

[0050] 本发明另提出一种数据接收终端,用以接收数据发送端的视频和音频数据并同步地播放,该数据发送端在所发送的视频数据和音频数据中分别加入视频时间戳和音频时间戳,其中该数据接收终端包括:

[0051] 获取接收的音频数据及相应的音频时间戳,并记录本地当前时钟的单元;

[0052] 获取接收的视频数据及相应的视频时间戳,形成完整的视频数据帧,并记录本地当前时钟的单元;

[0053] 依据相邻音频数据帧的音频时间戳及本地当前时钟计算所述音频抖动时间的单元;

[0054] 每隔预定时间将一帧音频数据送给音频解码器,如果音频数据因误码而被丢弃,则生成一静默帧送给音频解码器的单元;以及

[0055] 参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器的单元。

[0056] 在本发明的一实施例中,在参考当前正在播放的音频时间戳、待处理的视频数据帧的视频时间戳以及该音频抖动时间,以决定是否将此视频数据帧交给视频解码器的单元中,如果该待处理的视频帧的视频时间戳大于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则不将视频帧交给视频解码器;如果该待处理的视频帧的视频时间戳小于等于当前正在播放的音频时间戳和一个预定常量及该音频抖动时间的累加值,则将视频帧交给视频解码器。

[0057] 在本发明中,如果出现音频帧出错或没有按正确时序发送给音频解码器,则接收终端会主动生成一个静默帧交给音频解码器,这样可以平滑声音且不会出现有杂音。同时由于每个音频帧及视频帧都加上了时间戳作为同步信息,就可以处理由于发送端、网络、接收端的处理而引起音频视频不同步,这样就可以保证音频视频在允许范围内同步。因此,采用本发明可以改善音频质量,避免出现杂音,也避免因长时间通话而导致音频视频不同步。

### 附图说明

[0058] 为了让本发明的上述目的、特征和优点能更明显易懂,以下结合附图对本发明的具体实施方式作详细说明,其中:

[0059] 图 1 示出根据本发明一实施例的系统框图。

[0060] 图 2 示出根据本发明一实施例的系统操作流程图。

[0061] 图 3 示出根据本发明一实施例的数据包格式。

[0062] 图 4 示出根据本发明一实施例的发送端上行数据流程图。

[0063] 图 5 示出根据本发明一实施例的接收终端下行数据流程图。

### 具体实施方式

[0064] 图 1 示出根据本发明一实施例的系统框图。其中按照 3G-324M 标准,在 3G324M 协议栈 100 中采用 H. 223 协议栈 110 用作语音、视频、用户数据和控制数据的多路复用和分离;采用 H. 263 编解码器 120 用作视频编码,采用 AMR 编解码器 130 用作音频编码,同时还采用了 H. 245 协议栈 140。视频设备 150 可向 H. 263 编解码器 120 提供视频数据,语音设备 160 可向 AMR 编解码器 130 提供音频数据,经过编码后于 H. 223 协议栈 110 形成数据包向 3G 信道 170 发送。相应地,经由 3G 信道 170 接收的数据可经 H. 223 协议栈 110 处理后,再分别经过 H. 263 编解码器 120 和 AMR 编解码器 130 解码后送至视频设备 150 和语音设备 160 进行播放。其中,H. 223 协议栈 110 又可进一步分为复用层 (MUXLayer) 和适配层。

[0065] 系统的具体操作流程图请参照图 2 所示,请结合参照图 1 所示,在数据发送端的上行过程中,通过 H. 263 视频编码产生视频帧,并进一步形成多个 AL-PDU 1-n (其包结构请参见图 3 所示);同时通过 AMR 音频编码产生音频帧,并进一步形成一个 AL-PDU。视频帧和音频帧经过复用层复用后,发送到 3G 信道。在数据接收终端下行过程中,经 3G 信道接收的数据首先于复用层解复用,产生多个视频的 AL-PDU 1-n 以及一个音频的 AL-PDU,其中多个视频的 AL-PDU 1-n 形成完整的视频帧后,再经过 H. 263 视频解码后输出播放,而一个音频的 AL-PDU 直接形成音频帧,再经过 AMR 音频解码后输出播放。

[0066] 根据本发明的实施例,在图 2 所示流程中所发送的视频数据 AL-PDU 1-n 和音频数据 AL-PDU 中均加入时间戳。具体如图 3 所示,原有 AL-PDU 包含可选序列号 (optional

sequence number); :AL-PDU 有效载荷域 (AL-PDU payloadfield);CRC 校验域 (CRC Field)。本实施例是在一个原 AL-PDU 的头部基础之上再增加一个字段,即时间戳 (Time Stamp),占  $N$  两个字节 (octet),  $N = 1, 2, 3, \dots$ , 此字段的作用是为了同步音频视频。

[0067] 根据本发明的实施例,如果一个帧的视频数据是由多个 H. 223AL2 组成,那它们包头的时戳应该是一样的。例如图 2 所示的 AL-PDU 1-n 均具有相同的时戳。在一个实施例中,可以本端采样点数据的首比特传输时间为基准度量,并以此计算时戳,该数据包含但不限于音频数据,及视频数据。

[0068] 以 ITU 3G324M 协议来实现视频通话 (videophone) 为例,在发送端加入时戳的具体流程可参照图 4 所示,其中系统框图及基本操作流程已示出于图 1 和图 2,下面的流程主要描述时戳的加入有关的步骤:

[0069] 首先介绍视频处理步骤 S11-S14,具体如下:于步骤 S11,本端视频上行线程从 H. 263 编码器获取一帧数据,放入视频缓冲区中;紧接着于步骤 S12,在本地记录下当前的视频时戳  $T1(V)$ ;然后于步骤 S13,从视频缓冲区中取出一帧视频数据,并将此分割成  $n$  块,每一块就是一个 AL-SDU (即 1 个数据单元),其中 AL-SDU 构成图 3 所示 AL-PDU 数据包中的有效载荷;之后于步骤 S14,通过 H. 223 适配层处理,形成 AL-PDU 数据包,将步骤 S12 记录下的时戳  $T1(V)$  作为这  $n$  块包头的时戳字段。

[0070] 其次介绍音频处理步骤 S15-S18,具体如下:于步骤 S15,本端音频上行线程从 AMR 编码器获取一帧数据,放入音频缓冲区中;紧接着于步骤 S16,记录下当前时戳  $T1(A)$ ;然后于步骤 S17,从音频缓冲区中取出一帧音频数据,以此作为音频 AL-SDU (即 1 个数据单元);之后于步骤 S18,通过 H. 223 适配层处理,将步骤 S16 记录下的时戳  $T1(A)$  作为这包头的时戳字段。

[0071] 然后,于步骤 S19 经 H. 223 复用处理,将几个视频和音频数据流通过某种组合复成一个数据流后,经 3G 信道发送给远端。

[0072] 根据本发明的实施例,在接收终端,如果音频帧出错或者没有按正确时序发送给音频解码器进行解码,则接收终端会主动生成一个静默帧给音频解码器,而视频帧是依据音频帧来进行同步。图 5 示出根据本发明一实施例的接收终端下行数据流程图。其中系统框图及基本操作流程已示出于图 1 和图 2,数据接收终端的音频视频同步方法包括以下步骤:

[0073] 于步骤 S21,经 H. 223 解复用 (DEMUX) 处理,获得多个视频数据单元 AL-SDU 以及一个音频数据单元 AL-SDU 后,经过判断数据单元是否为音频数据 AL-SDU 的步骤 S22 分离音频逻辑信道中的音频数据及相应的时戳  $T2(A)$  以及视频逻辑信道中的视频数据及相应的时戳  $T2(V)$ 。

[0074] 对于分离的音频数据,于步骤 S23 判断其 AL-SDU 是否存在误码,若是,则于步骤 S24 丢弃此 AL-SDU,否则,于步骤 S25 将音频数据帧 (即一个 AL-SDU) 及音频时戳保存在音频帧缓冲区,之后于步骤 S26,同时记录下此时本地时钟  $T3(A)$ ,然后,于步骤 S27,依照下述公式计算音频抖动时间:

[0075] 
$$\text{jitter} = (T2(A)(n+1) - T2(A)(n)) - (T3(A)(n+1) - T3(A)(n)), n = 1, 2, 3, \dots$$
, 代表帧的序列号。

[0076] 对于分离的视频数据,首先于步骤 S28 放在 AL-SDU 缓冲区中,之后,于步骤 S29,以

图像开始码作为帧的分界标识来获取一个完整的视频数据帧,同时于步骤 S30 记录下此时本地时钟 T3(V),然后于步骤 S31,将视频数据帧及视频时间戳放入视频缓冲区中。

[0077] 在音频解码线程中,会于步骤 S32,每隔预定时间(按照 50 帧/秒的帧率可取 20ms)从音频缓存区取出一帧音频数据,然而如果于步骤 S33 发现缓存区 B(A) 为空(由于误码的音频帧已于先前的步骤 S24 被丢弃),则于步骤 S34 生成一个静默帧,否则,直接于步骤 S35 将音频帧送给音频解码器。

[0078] 静默帧作用主要有两个,一个是为了消除杂音,另一个作为是为了同步。这主要是考虑一些异常情况,比如音频帧因网络传输而导致丢失、延时、产生误码,因此当音频帧有一个固定的时序后,视频帧可以通过音频帧来同步。

[0079] 在视频解码线程中,会于步骤 S36,从视频帧缓冲区中获取视频帧,然后,于步骤 S37,视频帧交给视频解码器前要参考时间戳,具体地说,参考当前正在播放的音频时间戳 T4(A)、待处理的视频数据帧的视频时间戳 T4(V) 以及音频抖动时间 jitter,如果此视频帧的视频时间戳 T4(V) 大于当前正在播放的音频帧的音频时间戳 T4(A) 和一个常量 T 及音频抖动时间 jitter 的累加值,则此时不能将视频帧交给视频解码器,返回步骤 S37 继续判断;如果此视频帧的时间戳 T4(V) 小于当前正在播放的音频时间戳 T4(A) 和常量 T 及音频抖动时间 jitter 的累加值时,则此时可以于步骤 S38 将视频帧交给 H. 263 视频解码器。常量 T 考虑了发送端无法保证速率、误码等因素来选取,例如取 200ms。

[0080] 在低码率无线信道终端在高误码率环境,信道中存在误码是客观存在的,如果音频帧出错或没有接收到端没有按正确时序发送给音频解码器,会导致有杂音出现,给用户造成不良的使用体验。所以在本发明中,如果出现音频帧出错或没有按正确时序发送给音频解码器,则接收终端会主动生成一个静默帧交给音频解码器,这样可以平滑声音且不会出现有杂音。同时由于每个音频帧及视频帧都加上了时间戳作为同步信息,就可以处理由于发送端、网络、接收端的处理而引起音频视频不同步,这样就可以保证音频视频在允许范围内同步(假设视频的帧率为 5 ~ 15 帧,则同步误差为 180ms)。因此,采用本发明可以改善音频质量,避免出现杂音,也避免因长时间通话而导致音频视频不同步。

[0081] 虽然本发明已以较佳实施例揭示如上,然其并非用以限定本发明,任何本领域技术人员,在不脱离本发明的精神和范围内,当可作些许的修改和完善,因此本发明的保护范围当以权利要求书所界定的为准。

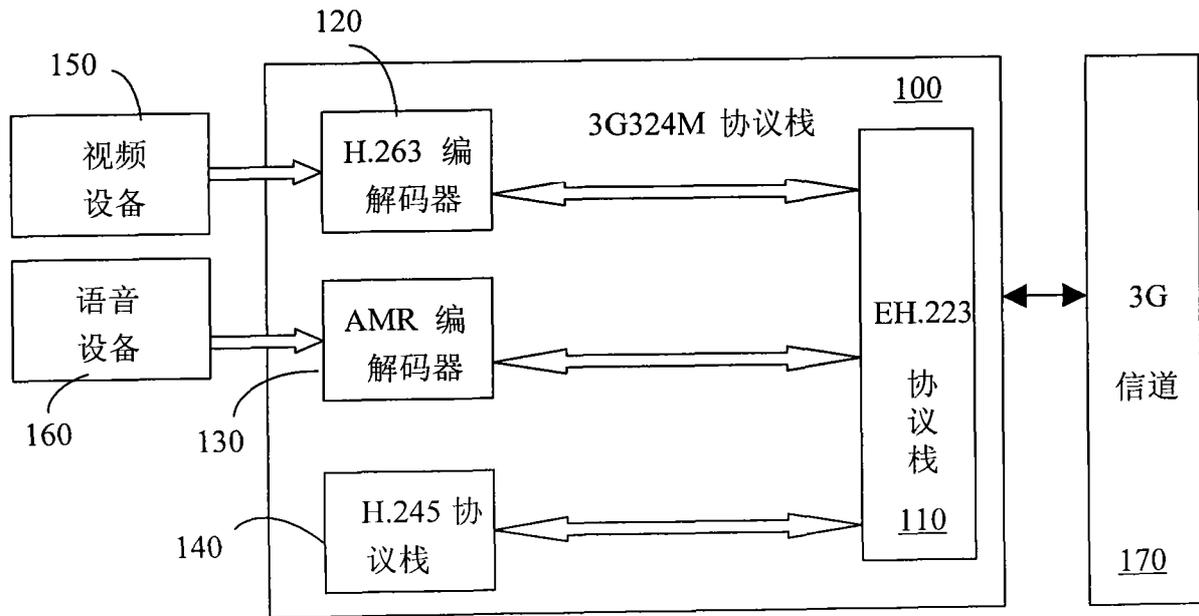


图 1

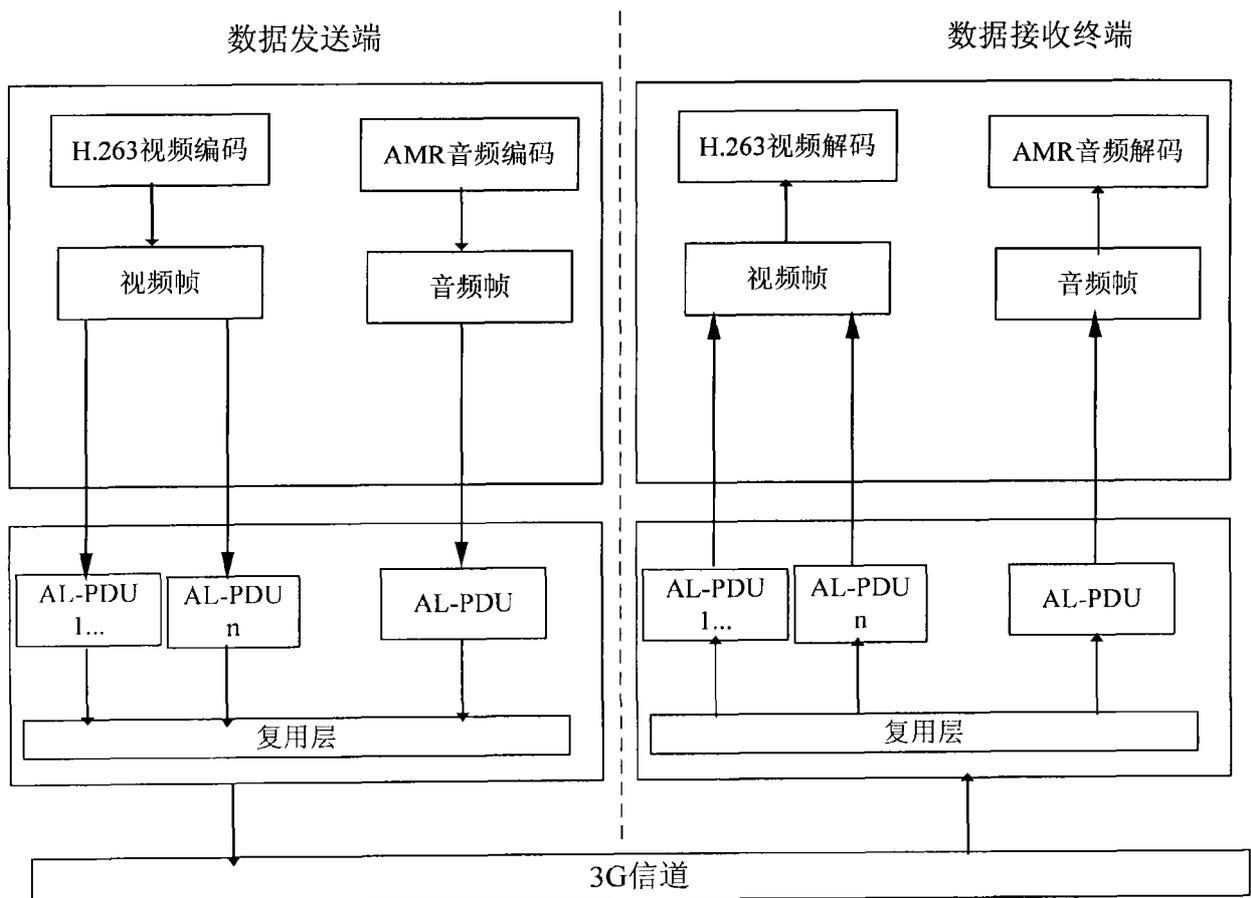


图 2

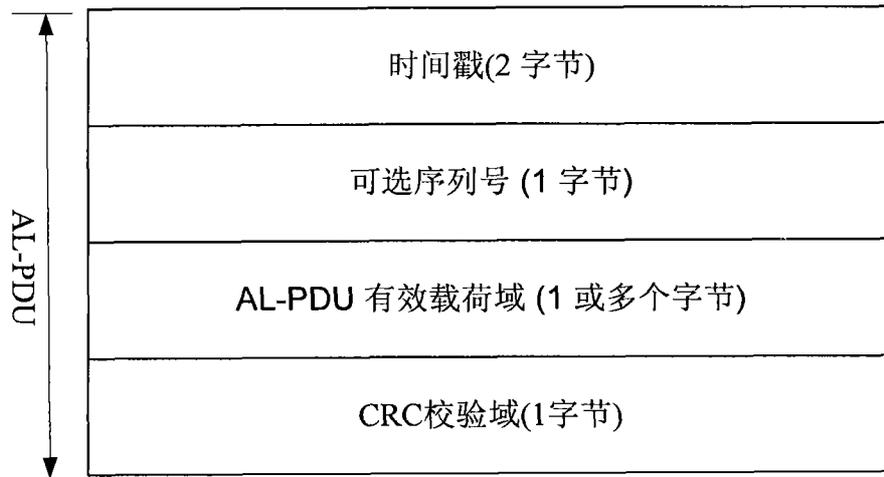


图 3

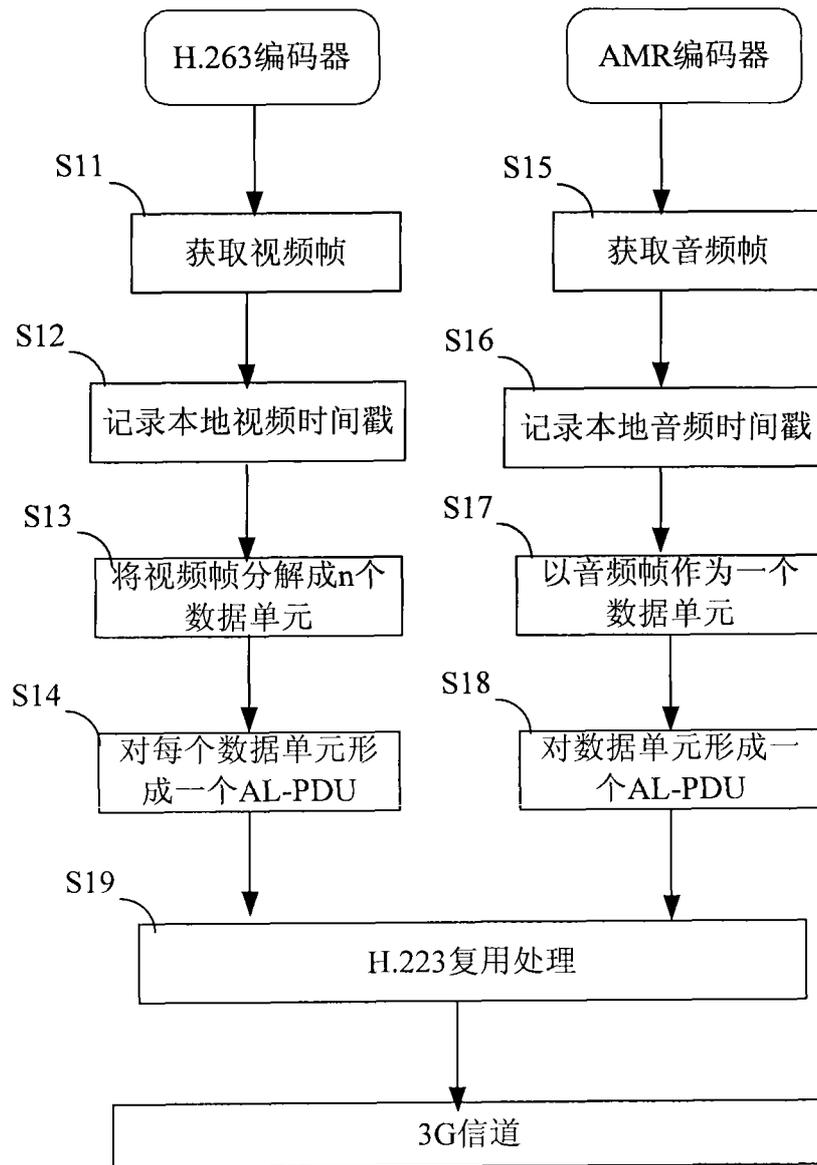


图 4

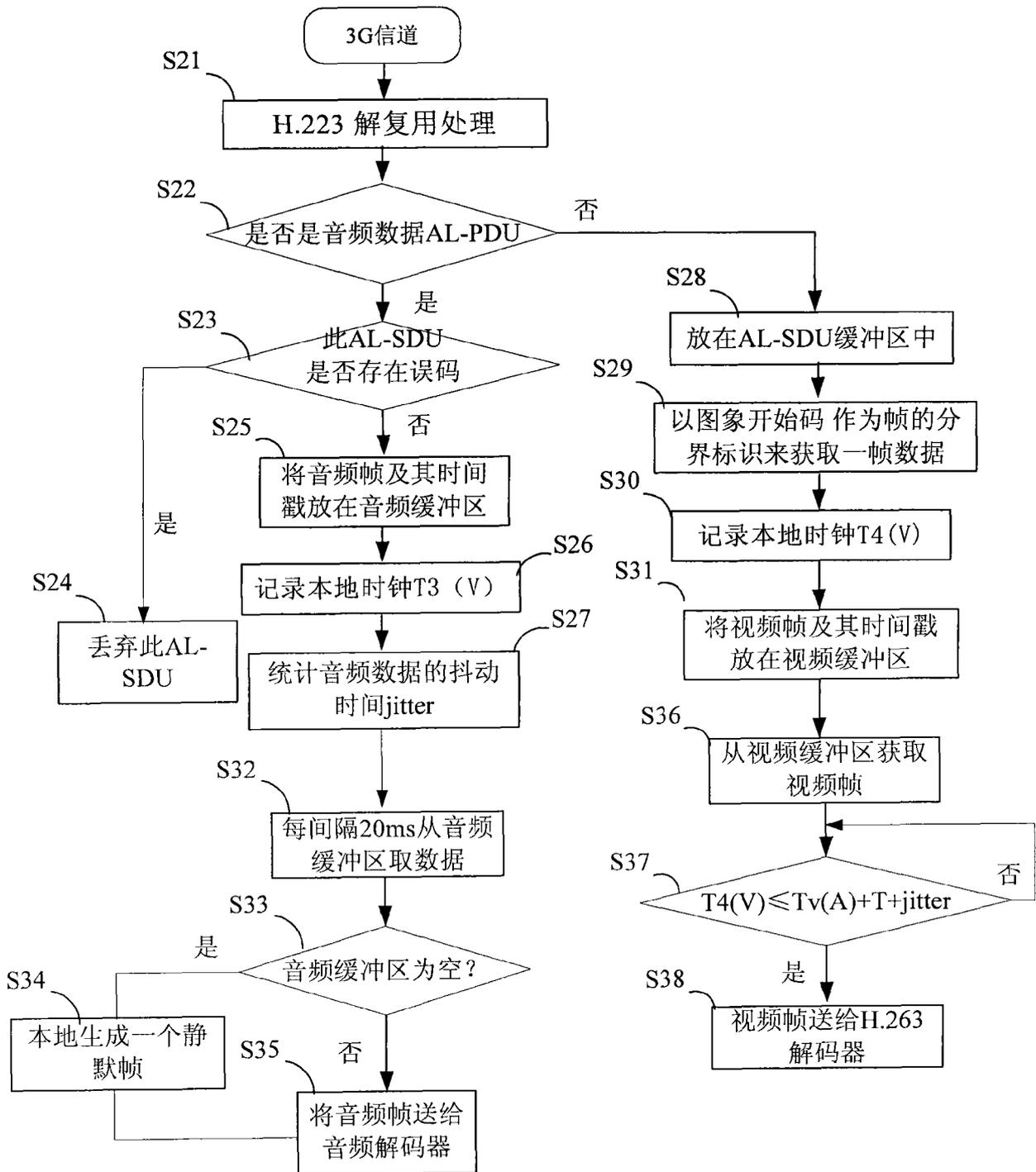


图 5