



(12)发明专利申请

(10)申请公布号 CN 107770597 A

(43)申请公布日 2018.03.06

(21)申请号 201710901655.X

(22)申请日 2017.09.28

(71)申请人 北京小鸟科技股份有限公司  
地址 102208 北京市昌平区回龙观龙冠商务中心5层501室

(72)发明人 刘红院 李厚鹏 董志松 孙立宁  
王自强 崔标 李伟 温泉  
董学明 杜键

(74)专利代理机构 北京瀚仁知识产权代理事务所(普通合伙) 11482  
代理人 郭文浩 杨晓莉

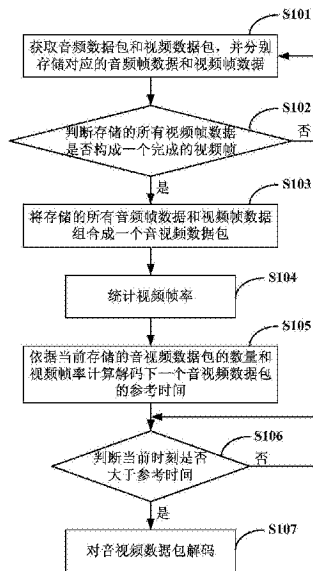
(51)Int. Cl.  
H04N 21/43(2011.01)  
H04N 21/44(2011.01)  
H04N 21/439(2011.01)

权利要求书3页 说明书10页 附图4页

(54)发明名称  
音视频同步方法及装置

(57)摘要

本发明涉及多媒体技术领域,具体提供了一种音视频同步方法及装置,旨在解决视频数据与音频数据不同步的技术问题。为此目的,本发明中的音视频同步方法包括下述步骤:获取音频数据包和视频数据包,并分别存储对应的音频帧数据和视频帧数据;将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包,并统计视频帧率;依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间;判断当前时刻是否大于参考时间:若是则对音视频数据包解码。同时,本发明中的装置可以执行上述方法的各步骤。本发明的技术方案,可以在网络传输带宽有限或网络环境复杂的情况下保证音视频播放具有较高的流畅性和同步性。



1. 一种音视频同步方法,其特征在于,所述方法包括:

获取音频数据包和视频数据包,并分别存储对应的音频帧数据和视频帧数据;

判断存储的所有视频帧数据是否构成一个完成的视频帧:若是则将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包,并统计视频帧率;

依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间;

判断当前时刻是否大于所述参考时间:若是则对所述音视频数据包解码。

2. 根据权利要求1所述的方法,其特征在于,“判断存储的所有视频帧数据是否构成一个完成的视频帧”的步骤具体包括:

依据当前获取的视频数据包中的视频数据,判断所述视频数据包是否为一帧视频的最后一个视频帧数据包,若是则存储的所有视频帧数据能够构成一个完成的视频帧。

3. 根据权利要求1所述的方法,其特征在于,“将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包”的步骤具体包括:

将视频帧数据和音频帧数据依次设置在预设的音视频包头之后,形成一个包含连续存储的所述预设的音视频包头、视频帧数据和音频帧数据的音视频数据包;

其中,所述预设的音视频包头包括视频帧数及每帧的长度,以及音频帧数及每帧的长度。

4. 根据权利要求1所述的方法,其特征在于,“统计视频帧率”的步骤具体包括按照下式所示的方法计算预设的统计时间间隔T内的视频帧率 $F_R$ :

$$F_R = \frac{F_N}{T}$$

其中,所述 $F_N$ 为预设的统计时间间隔T内的视频帧数。

5. 根据权利要求1-4任一项所述的方法,其特征在于,“依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间”的步骤具体包括:

按照各预设的数量阈值由小到大的顺序,依次对所述音视频数据包数量,以及所述各预设的数量阈值进行比较;

当比较得到第一个大于所述音视频数据包数量的数量阈值时,依据所得到的数量阈值,及预设的时间匹配表,匹配出该数据阈值对应的时间系数;所述预设的时间匹配表包括所述各预设的数据阈值,以及与各预设的数据阈值对应的时间系数;

当未比较得到大于所述音视频数据包数据的数据阈值时,选取预设的时间系数;

依据所匹配或所选取的时间系数,并按照下式计算下一个音视频数据包的参考时间:

$$T_{\text{reference}(n+1)} = T_{\text{reference}(n)} + T_{\text{frame}} \times k$$

其中,所述 $T_{\text{reference}(n+1)}$ 为第n+1个音视频数据包的参考时间,所述 $T_{\text{reference}(n)}$ 为第n个音视频数据包的参考时间, $n \geq 0$ ;所述k为时间系数, $k \geq 0$ ;所述 $T_{\text{frame}}$ 为视频帧的时间间隔,且 $T_{\text{frame}} = 1000/F_R$ , $F_R$ 为视频帧率。

6. 根据权利要求5所述的方法,其特征在于,“依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间”的步骤之前包括:

判断当前视频帧率与上一个视频帧率的偏差是否大于预设的误差值:若是则依据当前

视频帧率和所述的当前存储的音视频数据包的数量计算解码下一个音视频数据包的参考时间。

7. 根据权利要求1-4任一项所述的方法,其特征在于,“对所述音视频数据包解码”的步骤之前包括:

判断在音视频数据包解码通道中当前缓存的音频帧数是否大于预设缓存阈值:若是则清除所述音视频数据包解码通道中的音频数据,若否则对所述音视频数据包解码。

8. 一种音视频同步装置,其特征在于,所述方法包括:

第一接收模块,配置为获取音频数据包和视频数据包,并分别存储对应的音频帧数据和视频帧数据;

第二接收模块,配置为判断存储的所有视频帧数据是否构成一个完成的视频帧:若是则将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包,并统计视频帧率;

第一解码模块,配置为依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间;

第二解码模块,配置为判断当前时刻是否大于所述参考时间:若是则对所述音视频数据包解码。

9. 根据权利要求8所述的装置,其特征在于,所述第二接收模块包括判断单元;所述判断单元,配置为依据当前获取的视频数据包中的视频数据,判断所述视频数据包是否为一帧视频的最后一个视频帧数据包,若是则存储的所有视频帧数据能够构成一个完成的视频帧。

10. 根据权利要求8所述的装置,其特征在于,所述第二接收模块包括数据组合单元;所述数据组合单元,配置为将视频帧数据和音频帧数据依次设置在预设的音视频包头之后,形成一个包含连续存储的所述预设的音视频包头、视频帧数据和音频帧数据的音视频数据包;

其中,所述预设的音视频包头包括视频帧数及每帧的长度,以及音频帧数及每帧的长度。

11. 根据权利要求8所述的装置,其特征在于,所述第二接收模块包括统计单元;所述统计单元,配置为按照下式所示的方法计算预设的统计时间间隔T内的视频帧率 $F_R$ :

$$F_R = \frac{F_N}{T}$$

其中,所述 $F_N$ 为预设的统计时间间隔T内的视频帧数。

12. 根据权利要求8-11任一项所述的装置,其特征在于,所述第一解码模块包括比较单元和计算单元;

所述比较单元,配置为按照各预设的数量阈值由小到大的顺序,依次对所述音视频数据包数量,以及所述各预设的数量阈值进行比较;

当比较得到第一个大于所述音视频数据包数量的数量阈值时,依据所得到的数量阈值,及预设的时间匹配表,匹配出该数据阈值对应的时间系数;所述预设的时间匹配表包括所述各预设的数据阈值,以及与各预设的数据阈值对应的时间系数;

当未比较得到大于所述音视频数据包数据的数据阈值时,选取预设的时间系数;

所述计算单元,配置为依据所述比较单元所匹配或所选取出的时间系数,并按照下式计算下一个音视频数据包的参考时间:

$$T_{\text{reference}(n+1)} = T_{\text{reference}(n)} + T_{\text{frame}} \times k$$

其中,所述 $T_{\text{reference}(n+1)}$ 为第 $n+1$ 个音视频数据包的参考时间,所述 $T_{\text{reference}(n)}$ 为第 $n$ 个音视频数据包的参考时间, $n \geq 0$ ;所述 $k$ 为时间系数, $k \geq 0$ ;所述 $T_{\text{frame}}$ 为视频帧的时间间隔,且 $T_{\text{frame}} = 1000/F_R$ , $F_R$ 为视频帧率。

13. 根据权利要求8-11任一项所述的装置,其特征在于,所述装置还包括第一判断模块;所述第一判断模块,配置为判断当前视频帧率与上一个视频帧率的偏差是否大于预设的误差值;若是则依据当前视频帧率和所述的当前存储的音视频数据包的数目计算解码下一个音视频数据包的参考时间。

14. 根据权利要求8-11任一项所述的装置,其特征在于,所述装置还包括第二判断模块;所述第二判断模块,配置为判断在音视频数据包解码通道中当前缓存的音频帧数是否大于预设缓存阈值;若是则清除所述音视频数据包解码通道中的音频数据,若否则对所述音视频数据包解码。

15. 一种存储装置,其中存储有多条程序,其特征在于,所述程序适于由处理器加载并执行以实现权利要求1-7任一项所述的音视频同步方法。

16. 一种处理装置,包括

处理器,适于执行各条程序;以及

存储设备,适于存储多条程序;

其特征在于,所述程序适于由处理器加载并执行以实现:

权利要求1-7任一项所述的音视频同步方法。

17. 一种音视频解码装置,其特征在于,所述装置包括视频显示模块、音频播放模块和权利要求8-14任一项所述的音视频同步装置;

所述视频显示模块,配置为显示经所述音视频同步装置解码后的视频数据;

所述音频播放模块,配置为播放经所述音视频同步装置解码后的音频数据。

## 音视频同步方法及装置

### 技术领域

[0001] 本发明涉及多媒体技术领域,具体涉及一种音视频同步方法及装置。

### 背景技术

[0002] 目前,随着高清化和IP化技术的发展,越来越多的视频数据和音频数据通过网络进行传输。具体地,通过视频采集设备对采集到的实时视频数据和实时音频数据进行编码,并将编码后的数据发送至解码设备进行解码。但是,在上述传输过程中,会存在网络传输带宽有限或网络环境复杂的情况,导致视频数据和音频数据不同步,严重影响音视频的观看效果。

### 发明内容

[0003] 为了解决现有技术中的上述问题,即为了解决视频数据与音频数据不同步的技术问题,本发明提供了一种音视频同步方法及装置。

[0004] 在第一方面,本发明中音视频同步方法包括:

[0005] 获取音频数据包和视频数据包,并分别存储对应的音频帧数据和视频帧数据;

[0006] 判断存储的所有视频帧数据是否构成一个完成的视频帧:若是则将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包,并统计视频帧率;

[0007] 依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间;

[0008] 判断当前时刻是否大于所述参考时间:若是则对所述音视频数据包解码。

[0009] 进一步地,本发明提供的一个优选技术方案为:

[0010] “判断存储的所有视频帧数据是否构成一个完成的视频帧”的步骤具体包括:

[0011] 依据当前获取的视频数据包中的视频数据,判断所述视频数据包是否为一帧视频的最后一个视频帧数据包,若是则存储的所有视频帧数据能够构成一个完成的视频帧。

[0012] 进一步地,本发明提供的一个优选技术方案为:

[0013] “将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包”的步骤具体包括:

[0014] 将视频帧数据和音频帧数据依次设置在预设的音视频包头之后,形成一个包含连续存储的所述预设的音视频包头、视频帧数据和音频帧数据的音视频数据包;

[0015] 其中,所述预设的音视频包头包括视频帧数及每帧的长度,以及音频帧数及每帧的长度。

[0016] 进一步地,本发明提供的一个优选技术方案为:

[0017] “统计视频帧率”的步骤具体包括按照下式所示的方法计算预设的统计时间间隔T内的视频帧率 $F_R$ :

$$F_R = \frac{F_N}{T}$$

[0018] 其中,所述 $F_N$ 为预设的统计时间间隔 $T$ 内的视频帧数。

[0019] 进一步地,本发明提供的一个优选技术方案为:

[0020] “依据当前存储的音视频数据包数量和视频帧率计算解码下一个音视频数据包的参考时间”的步骤具体包括:

[0021] 按照各预设的数量阈值由小到大的顺序,依次对所述音视频数据包数量,以及所述各预设的数量阈值进行比较;

[0022] 当比较得到第一个大于所述音视频数据包数量的数量阈值时,依据所得到的数量阈值,及预设的时间匹配表,匹配出该数据阈值对应的时间系数;所述预设的时间匹配表包括所述各预设的数据阈值,以及与各预设的数据阈值对应的时间系数;

[0023] 当未比较得到大于所述音视频数据包数据的数据阈值时,选取预设的时间系数;

[0024] 依据所匹配或所选取的时间系数,并按照下式计算下一个音视频数据包的参考时间:

$$T_{\text{reference}(n+1)} = T_{\text{reference}(n)} + T_{\text{frame}} \times k$$

[0025] 其中,所述 $T_{\text{reference}(n+1)}$ 为第 $n+1$ 个音视频数据包的参考时间,所述 $T_{\text{reference}(n)}$ 为第 $n$ 个音视频数据包的参考时间, $n \geq 0$ ;所述 $k$ 为时间系数, $k \geq 0$ ;所述 $T_{\text{frame}}$ 为视频帧的时间间隔,且 $T_{\text{frame}} = 1000/F_R$ , $F_R$ 为视频帧率。

[0026] 进一步地,本发明提供的一个优选技术方案为:

[0027] “依据当前存储的音视频数据包数量和视频帧率计算解码下一个音视频数据包的参考时间”的步骤之前包括:

[0028] 判断当前视频帧率与上一个视频帧率的偏差是否大于预设的误差值:若是则依据当前视频帧率和所述的当前存储的音视频数据包的数量计算解码下一个音视频数据包的参考时间。

[0029] 进一步地,本发明提供的一个优选技术方案为:

[0030] “对所述音视频数据包解码”的步骤之前包括:

[0031] 判断在音视频数据包解码通道中当前缓存的音频帧数是否大于预设缓存阈值:若是则清除所述音视频数据包解码通道中的音频数据,若否则对所述音视频数据包解码。

[0032] 在第二方面,本发明中音视频同步装置包括:

[0033] 第一接收模块,配置为获取音频数据包和视频数据包,并分别存储对应的音频帧数据和视频帧数据;

[0034] 第二接收模块,配置为判断存储的所有视频帧数据是否构成一个完成的视频帧:若是则将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包,并统计视频帧率;

[0035] 第一解码模块,配置为依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间;

[0036] 第二解码模块,配置为判断当前时刻是否大于所述参考时间:若是则对所述音视频数据包解码。

[0037] 进一步地,本发明提供的一个优选技术方案为:

[0038] 所述第二接收模块包括判断单元;所述判断单元,配置为依据当前获取的视频数据包中的视频数据,判断所述视频数据包是否为一帧视频的最后一个视频帧数据包,若是则存储的所有视频帧数据能够构成一个完成的视频帧。

[0039] 进一步地,本发明提供的一个优选技术方案为:

[0040] 所述第二接收模块包括数据组合单元;所述数据组合单元,配置为将视频帧数据和音频帧数据依次设置在预设的音视频包头之后,形成一个包含连续存储的所述预设的音视频包头、视频帧数据和音频帧数据的音视频数据包;

[0041] 其中,所述预设的音视频包头包括视频帧数及每帧的长度,以及音频帧数及每帧的长度。

[0042] 进一步地,本发明提供的一个优选技术方案为:

[0043] 所述第二接收模块包括统计单元;所述统计单元,配置为按照下式所示的方法计算预设的统计时间间隔T内的视频帧率 $F_R$ :

$$F_R = \frac{F_N}{T}$$

[0044] 其中,所述 $F_N$ 为预设的统计时间间隔T内的视频帧数。

[0045] 进一步地,本发明提供的一个优选技术方案为:

[0046] 所述第一解码模块包括比较单元和计算单元;

[0047] 所述比较单元,配置为按照各预设的数量阈值由小到大的顺序,依次对所述音视频数据包数量,以及所述各预设的数量阈值进行比较;

[0048] 当比较得到第一个大于所述音视频数据包数量的数量阈值时,依据所得到的数量阈值,及预设的时间匹配表,匹配出该数据阈值对应的时间系数;所述预设的时间匹配表包括所述各预设的数据阈值,以及与各预设的数据阈值对应的时间系数;

[0049] 当未比较得到大于所述音视频数据包数据的数据阈值时,选取预设的时间系数;

[0050] 所述计算单元,配置为依据所述比较单元所匹配或所选取出的时间系数,并按照下式计算下一个音视频数据包的参考时间:

$$T_{\text{reference}(n+1)} = T_{\text{reference}(n)} + T_{\text{frame}} \times k$$

[0051] 其中,所述 $T_{\text{reference}(n+1)}$ 为第n+1个音视频数据包的参考时间,所述 $T_{\text{reference}(n)}$ 为第n个音视频数据包的参考时间, $n \geq 0$ ;所述k为时间系数, $k \geq 0$ ;所述 $T_{\text{frame}}$ 为视频帧的时间间隔,且 $T_{\text{frame}} = 1000/F_R$ , $F_R$ 为视频帧率。

[0052] 进一步地,本发明提供的一个优选技术方案为:

[0053] 所述装置还包括第一判断模块;所述第一判断模块,配置为判断当前视频帧率与上一个视频帧率的偏差是否大于预设的误差值;若是则依据当前视频帧率和所述的当前存储的音视频数据包的数量计算解码下一个音视频数据包的参考时间。

[0054] 进一步地,本发明提供的一个优选技术方案为:

[0055] 所述装置还包括第二判断模块;所述第二判断模块,配置为判断在音视频数据包解码通道中当前缓存的音频帧数是否大于预设缓存阈值;若是则清除所述音视频数据包解码通道中的音频数据,若否则对所述音视频数据包解码。

[0056] 在第三方面,本发明中存储装置存储有多条程序,所述程序适于由处理器加载并执行以上述技术方案所述的音视频同步方法。

[0057] 在第四方面,本发明中处理装置包括:

[0058] 处理器,适于执行各条程序;以及

[0059] 存储设备,适于存储多条程序;

[0060] 所述程序适于由处理器加载并执行以实现上述技术方案所述的音视频同步方法。

[0061] 在第五方面,本发明中音视频解码装置包括视频显示模块、音频播放模块和上述音视频同步装置所述的音视频同步装置;

[0062] 所述视频显示模块,配置为显示经所述音视频同步装置解码后的视频数据;

[0063] 所述音频播放模块,配置为播放经所述音视频同步装置解码后的音频数据。

[0064] 与最接近的现有技术相比,上述技术方案至少具有以下有益效果:

[0065] 1、本发明的音视频同步方法中在当前视频帧率与上一个视频帧率的偏差大于预设的误差值时,可以依据当前视频帧率和当前存储的音视频数据包的数量计算解码下一个音视频数据包的参考时间,使得音视频发送终端发送的视频帧率发生改变后,及时调整参考时间,保证音视频播放的流畅性和同步性。

[0066] 2、本发明的音视频同步方法中在判断当前时刻是否大于参考时间时时,采用执行音视频同步方法的设备的运行时间与参考时间进行比较,可以保证音视频同步过程中时间的一致性。

[0067] 3、本发明的音视频同步方法中在音视频数据包解码通道中当前缓存的音频帧数大于预设缓存阈值时,可以在清除音视频数据包解码通道中的音频数据后再解码音视频数据包,避免音视频数据包解码通道中因缓存数据较多导致音视频不同步的问题发生。

## 附图说明

[0068] 图1是本发明实施例中音视频同步方法的主要步骤流程图;

[0069] 图2是本发明实施例中音视频同步装置的结构示意图;

[0070] 图3是本发明实施例中一个音视频数据包的示意图;

[0071] 图4是本发明实施例中获取音频/视频数据的主要步骤流程图;

[0072] 图5是本发明实施例中解码音频/视频数据的主要步骤流程图。

## 具体实施方式

[0073] 下面参照附图来描述本发明的优选实施方式。本领域技术人员应当理解的是,这些实施方式仅仅用于解释本发明的技术原理,并非旨在限制本发明的保护范围。

[0074] 目前,视频采集设备在分别接收到视频数据和音频数据后,对所接收的视频数据和音频数据进行编码,并将编码后的数据发送至解码设备进行解码播放。在此过程中,可能会由于网络传输带宽有限或网络环境复杂等问题,导致播放的音视频不同步。基于此,本发明提供了一种音视频同步方法,可以保证播放的音视频具有较高的流畅性和同步性。

[0075] 下面结合附图,对本发明提供的音视频同步方法进行说明。

[0076] 参阅附图1,图1示例性示出了本实施例中音视频同步方法的主要步骤。如图1所示,本实施例中音视频同步方法包括下述步骤:



[0077] 步骤S101:获取音频数据包和视频数据包,并分别存储对应的音频帧数据和视频帧数据。

[0078] 具体地,本实施例中在获取到音频数据包后,首先解析该音频数据包的包头,然后将实际的音频帧数据存储至预先设置的音频缓存区中。类似地,在获取到视频数据包后,首先解析该视频数据包的包头,然后将实际的视频帧数据存储至预先设置的视频缓存区中。

[0079] 步骤S102:判断存储的所有视频帧数据是否构成一个完成的视频帧,若是则执行步骤S103,若否则返回步骤S101。

[0080] 具体地,本实施例中可以按照下述步骤判断视频缓存区中存储的所有视频帧数据是否构成一个完成的视频帧:

[0081] 依据当前获取的视频数据包中的视频数据,判断视频数据包是否为一帧视频的最后一个视频帧数据包,若是则视频缓存区中存储的所有视频帧数据能够构成一个完成的视频帧。例如,在解析视频数据包的包头后,依据包头信息中的帧标识符判断当前获取的视频数据包是否为一帧视频的最后一个视频帧数据包。

[0082] 步骤S103:将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包。具体地,本实施例中可以按照下述步骤将音频缓存区中存储的所有音频帧数据和视频缓存区中存储的视频帧数据组合成一个音视频数据包:

[0083] 将视频帧数据和音频帧数据依次设置在预设的音视频包头之后,形成一个包含连续存储的预设的音视频包头、视频帧数据和音频帧数据的音视频数据包。其中,本实施例中音视频数据包包括一帧视频帧数据和多帧音频帧数据。同时,预设的音视频包头包括视频帧数及每帧的长度,以及音频帧数及每帧的长度。

[0084] 参阅附图3,图3示例性示出了本实施例中一个音视频数据包。如图3所示,本实施例中音视频数据包包括包头、一帧视频帧数据和两帧音频帧数据。其中,包头中视频帧为1,音频帧为2,一帧视频的长度为2字节,一帧音频的长度为2字节。

[0085] 步骤S104:统计视频帧率。

[0086] 具体地,本实施例中可以按照下式(1)所示的方法计算预设的统计时间间隔T内的视频帧率 $F_R$ :

$$F_R = \frac{F_N}{T} \quad (1)$$

[0087] 公式(1)中各参数 $F_N$ 的含义为预设的统计时间间隔T内的视频帧数。

[0088] 在本实施例的一个优选实施方案中,可以将预设的统计时间间隔T设置为一个较小值,如2s,从而提高视频帧率 $F_R$ 的实时性。

[0089] 步骤S105:依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间。

[0090] 具体地,本实施例中参考时间的初始值为0,同时可以按照下述步骤计算解码下一个音视频数据包的参考时间:

[0091] 步骤S1051:按照各预设的数量阈值由小到大的顺序,依次对音视频数据包数量,以及各预设的数量阈值进行比较。

[0092] 具体地,本实施例中数量阈值的个数可以依据实际工况设定,如数据阈值可以为

三个。

[0093] 步骤S1052:当比较得到第一个大于音视频数据包数量的数量阈值时,依据所得到的数量阈值,及预设的时间匹配表,匹配出该数据阈值对应的的时间系数。其中,预设的时间匹配表包括各预设的数据阈值,以及与各预设的数据阈值对应的的时间系数。

[0094] 例如,本实施例中预先设置3个数据阈值 $THR1=5$ 、 $THR2=20$ 和 $THR3=45$ ,且数据阈值 $THR1$ 、 $THR2$ 和 $THR3$ 对应的的时间系数分别为2、1和1/2。当音视频数据包的数据为15,按照步骤S1051的方法可以得到第一个大于音视频数据包数量的数量阈值为 $THR2$ ,匹配出的时间系数为1。

[0095] 进一步地,本实施例中步骤S1052还包括:当未比较得到大于音视频数据包数据的数据阈值时,选取预设的时间系数。例如,该预设的时间系数可以为1/4。

[0096] 步骤S1053:依据所匹配或所选取的时间系数,并按照下式(2)计算下一个音视频数据包的参考时间:

$$T_{\text{reference}(n+1)} = T_{\text{reference}(n)} + T_{\text{frame}} \times k \quad (2)$$

[0097] 公式(2)中各参数含义为:

[0098]  $T_{\text{reference}(n+1)}$ 为第 $n+1$ 个音视频数据包的参考时间,即下一个音视频数据包的参考时间。 $T_{\text{reference}(n)}$ 为第 $n$ 个音视频数据包的参考时间, $n \geq 0$ ,即当前音视频数据包的参考时间。 $k$ 为步骤S1052中匹配或选取出的时间系数, $k \geq 0$ ;  $T_{\text{frame}}$ 为视频帧的时间间隔,且 $T_{\text{frame}} = 1000/F_r$ 。

[0099] 进一步地,本实施例的一个优选实施方案中,图1所示的步骤S105之前还包括下述步骤:

[0100] 判断当前视频帧率与上一个视频帧率的偏差是否大于预设的误差值:若是则依据当前视频帧率和所述的当前存储的音视频数据包的数量计算解码下一个音视频数据包的参考时间。

[0101] 本实施例中在前视频帧率与上一个视频帧率的偏差大于预设的误差值时,判断为视频发送设备的视频帧率发生了改变,在此情况下需要及时依据当前视频帧率和当前存储的音视频数据包的数量计算解码下一个音视频数据包的参考时间,以保证解码的流畅性和音视频的同步性,即可以在网络传输过程中发生网络拥塞等异常状态时也可以保证解码的流畅性和音视频的同步性。

[0102] 步骤S106:判断当前时刻是否大于参考时间,若是则执行步骤S107,若否则返回步骤S106。

[0103] 具体地,本实施例中当前时刻指的是执行上述步骤的设备的运行时间,该运行时间从设备启动后一直递增,可以保证音视频同步过程中时间的一致性。

[0104] 步骤S107:对音视频数据包解码。

[0105] 具体地,由前述步骤S103可知,音视频数据包包括连续存储的包头、视频帧数据和音频帧数据,因此在对音视频数据包解码时限解码音频帧数据再解码视频帧数据。

[0106] 进一步地,本实施例的一个优选实施方案中,图1所示的步骤S107之前还包括下述步骤:

[0107] 判断在音视频数据包解码通道中当前缓存的音频帧数是否大于预设缓存阈值:若是则清除音视频数据包解码通道中的音频数据,若否则对音视频数据包解码。

[0108] 本实施例中通过阈值判断的方式清除音视频数据包解码通道中的音频数据,可以防止在对音视频数据解码较慢时,由于音视频数据包解码通道中缓存数据较多,导致音视频不能同步的情况发生。

[0109] 进一步地,在本实施例的一个优选实施方案中,通过并行设置数据获取与数据解码两个实施流程,使得该音视频同步方法可以同时获取音频/视频数据,以及解码音频/视频数据,提高音视频的解码效率。

[0110] 下面结合附图,对本实施例中数据获取与数据解码两个实施流程分别进行说明。

[0111] 参阅附图4,图4示例性示出了本实施例中获取音频/视频数据的主要步骤。如图4所示,本实施例中可以按照下述步骤获取音频/视频数据:

[0112] 步骤S201:与音视频发送终端建立通信连接。

[0113] 具体地,本实施例中音视频发送终端可以为监控相机、音视频编码设备或网络流媒体平台。

[0114] 步骤S202:接收视频发送终端发送的音频数据包和视频数据包,并分别存储对应的音频帧数据和视频帧数据。

[0115] 步骤S203:当存储的所有视频帧数据能构成一个完成的视频帧时,存储的所有音频帧数据和视频帧数据组合成一个音视频数据包,并将该音视频数据包放入解码缓存队列。

[0116] 步骤S204:统计视频帧率,并在统计完成后返回步骤S201。

[0117] 继续参阅附图5,图5示例性示出了本实施例中解码音频/视频数据的主要步骤。如图5所示,本实施例中可以按照下述步骤解码音频/视频数据:

[0118] 步骤S301:对参考时间进行初始化,并设置预设缓存阈值。

[0119] 步骤S302:依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间。

[0120] 步骤S303:判断当前时刻是否大于参考时间:若是则执行步骤S304,若否则返回步骤S303。

[0121] 步骤S304:判断在音视频数据包解码通道中当前缓存的音频帧数是否大于预设缓存阈值:若是则执行步骤S305,若否则执行步骤S306。

[0122] 步骤S305:清除音视频数据包解码通道中的音频数据,并在清除音频数据后执行步骤S306。

[0123] 步骤S306:对音视频数据包中的音频帧数据进行解码,并在解码后执行步骤S307。

[0124] 步骤S307:对音视频数据包中的视音频帧数据进行解码,并在解码后返回步骤S302。

[0125] 上述实施例中虽然将各个步骤按照上述先后次序的方式进行了描述,但是本领域技术人员可以理解,为了实现本实施例的效果,不同的步骤之间不必按照这样的次序执行,其可以同时(并行)执行或以颠倒的次序执行,这些简单的变化都在本发明的保护范围之内。

[0126] 基于与方法实施例相同的技术构思,本发明实施例还提供一种音视频同步装置。下面结合附图对该音视频同步装置进行具体说明。

[0127] 参阅附图2,图2示例性示出了本实施例中音视频同步装置的结构。如图2所示,本

实施例中音视频同步装置主要包括第一接收模块11、第二接收模块12、第一解码模块13和第二解码模块14。其中，第一接收模块11可以配置为获取音频数据包和视频数据包，并分别存储对应的音频帧数据和视频帧数据。第二接收模块12可以配置为判断存储的所有视频帧数据是否构成一个完成的视频帧；若是则将存储的所有音频帧数据和视频帧数据组合成一个音视频数据包，并统计视频帧率。第一解码模块13可以配置为依据当前存储的音视频数据包的数量和视频帧率计算解码下一个音视频数据包的参考时间。第二解码模块14可以配置为判断当前时刻是否大于参考时间；若是则对音视频数据包解码。

[0128] 具体地，本实施例中第一接收模块11与音视频发送终端建立通信连接后，获取其放的音频数据包和视频数据包。其中，第一接收模块11可以采用网络通信方式，如基于TCP协议的通信方式或基于UDP协议的通信方式，也可以采用基于RTSP协议、ONVIF协议或RTMP协议的通信方式。

[0129] 进一步地，本实施例中第二接收模块可以包括判断单元、数据组合单元和统计单元。

[0130] 具体地，本实施例中判断单元可以配置为依据当前获取的视频数据包中的视频数据，判断视频数据包是否为一帧视频的最后一个视频帧数据包，若是则存储的所有视频帧数据能够构成一个完成的视频帧。

[0131] 本实施例中数据组合单元可以配置为将视频帧数据和音频帧数据依次设置在预设的音视频包头之后，形成一个包含连续存储的预设的音视频包头、视频帧数据和音频帧数据的音视频数据包。

[0132] 本实施例中统计单元，配置为按照公式(1)所示的方法计算预设的统计时间间隔T内的视频帧率 $F_R$ 。

[0133] 进一步地，本实施例中第一解码模块可以包括比较单元和计算单元。

[0134] 具体地，本实施例中比较单元可以配置为按照各预设的数量阈值由小到大的顺序，依次对音视频数据包数量，以及各预设的数量阈值进行比较。其中，当比较得到第一个大于音视频数据包数量的数量阈值时，依据所得到的数量阈值，及预设的时间匹配表，匹配出该数据阈值对应的时间系数。当未比较得到大于音视频数据包数据的数据阈值时，选取预设的时间系数。

[0135] 本实施例中计算单元可以配置为依据比较单元所匹配或所选取出的时间系数，并按照公式(2)计算下一个音视频数据包的参考时间。

[0136] 进一步地，本实施例中音视频同步装置还可以包括第一判断模块。该第一判断模块可以配置为判断当前视频帧率与上一个视频帧率的偏差是否大于预设的误差值；若是则依据当前视频帧率和当前存储的音视频数据包的数量计算解码下一个音视频数据包的参考时间。

[0137] 进一步地，本实施例中音视频同步装置还可以包括第二判断模块。该第二判断模块可以配置为判断在音视频数据包解码通道中当前缓存的音频帧数是否大于预设缓存阈值；若是则清除音视频数据包解码通道中的音频数据，若否则对音视频数据包解码。

[0138] 上述音视频同步装置实施例可以用于执行上述音视频同步方法实施例，其技术原理、所解决的技术问题及产生的技术效果相似，所属技术领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的音视频同步的具体工作过程及有关说明，可以参考前

述音视频同步方法实施例中的对应过程,在此不再赘述。

[0139] 本领域技术人员可以理解,上述音视频同步装置还包括一些其他公知结构,例如处理器、控制器、存储器等,其中,存储器包括但不限于随机存储器、闪存、只读存储器、可编程只读存储器、易失性存储器、非易失性存储器、串行存储器、并行存储器或寄存器等,处理器包括但不限于CPLD/FPGA、DSP、ARM处理器、MIPS处理器等,为了不必要地模糊本公开的实施例,这些公知的结构未在图2中示出。

[0140] 应该理解,图2中的各个模块的数量仅仅是示意性的。根据实际需要,各模块可以具有任意的数量。

[0141] 进一步地,基于上述音视频同步方法实施例,本发明还提供了一种存储装置,该存储装置中存储有多条程序,且这些程序可以适于由处理器加载并执行以实现上述方法实施例所述的音视频同步方法。

[0142] 进一步地,基于上述音视频同步方法实施例,本发明还提供了一种处理装置,该处理装置可以包括处理器和存储设备。其中,处理器可以适于执行各条程序,存储设备可以适于存储多条程序,且这些程序可以适于由处理器加载并执行以实现上述方法实施例所述的音视频同步方法。

[0143] 进一步地,基于上述音视频同步装置实施例,本发明还提供了一种音视频解码装置,该装置可以包括视频显示模块、音频播放模块和上述装置实施例所述的音视频同步装置。其中,视频显示模块可以配置为显示经音视频同步装置解码后的视频数据。音频播放模块可以配置为播放经音视频同步装置解码后的音频数据。

[0144] 本领域技术人员可以理解,可以对实施例中的装置中的模块进行自适应性地改变并且把它们设置在与该实施例不同的一个或多个装置中。可以把实施例中的模块或单元或组件组合成一个模块或单元或组件,以及此外可以把它分成多个子模块或子单元或子组件。除了这样的特征和/或过程或者单元中的至少一些是相互排斥之外,可以采用任何组合对本说明书(包括伴随的权利要求、摘要和附图)中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述,本说明书(包括伴随的权利要求、摘要和附图)中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

[0145] 此外,本领域的技术人员能够理解,尽管在此所述的一些实施例包括其它实施例中所包括的某些特征而不是其它特征,但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如,在本发明的权利要求书中,所要求保护的实施例的任意之一都可以以任意的组合方式来使用。

[0146] 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制,并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中,不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的PC来实现。在列举了若干装置的单元权利要求中,这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

[0147] 至此,已经结合附图所示的优选实施方式描述了本发明的技术方案,但是,本领域技术人员容易理解的是,本发明的保护范围显然不局限于这些具体实施方式。在不偏离本

发明的原理的前提下,本领域技术人员可以对相关技术特征作出等同的更改或替换,这些更改或替换之后的技术方案都将落入本发明的保护范围之内。

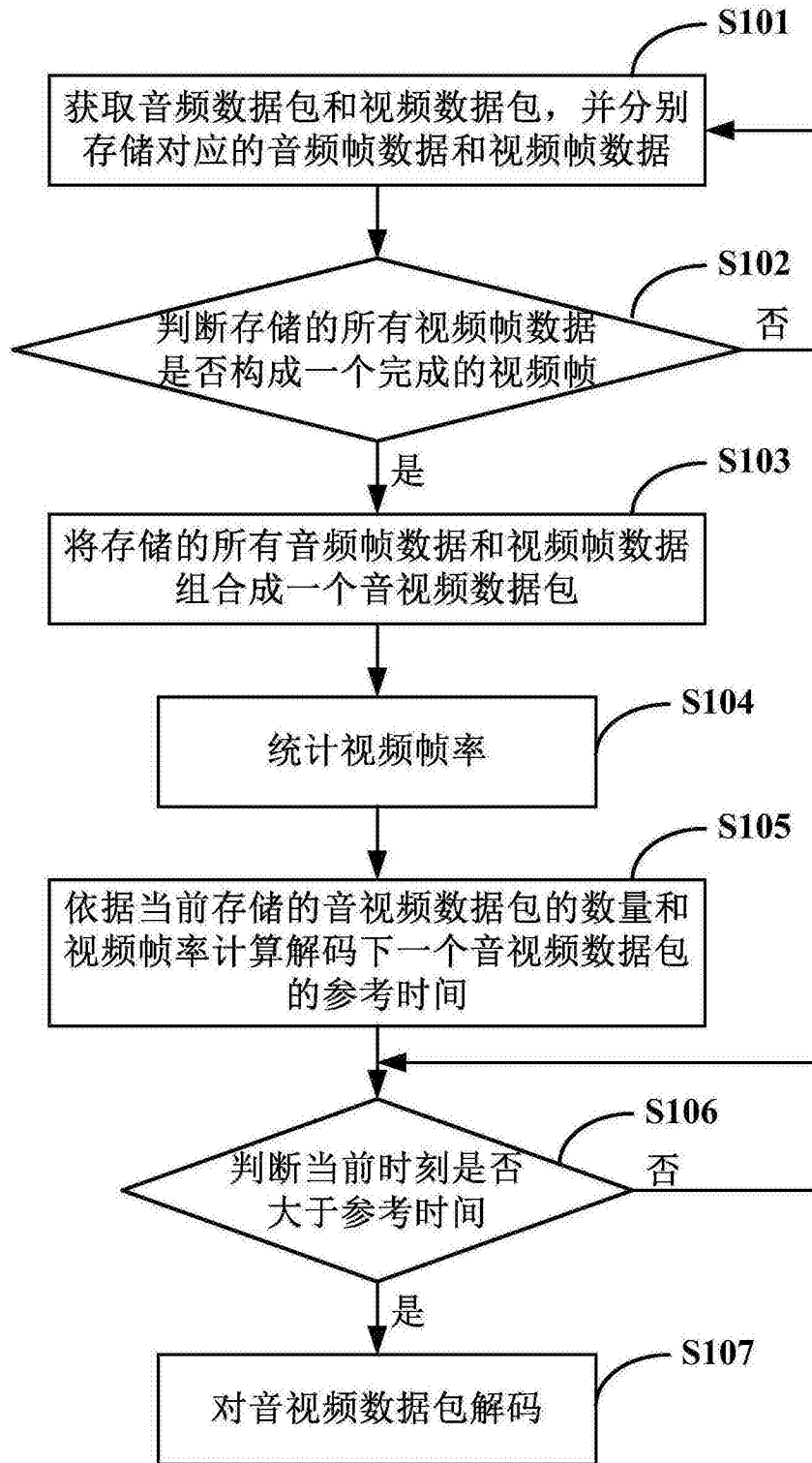


图1

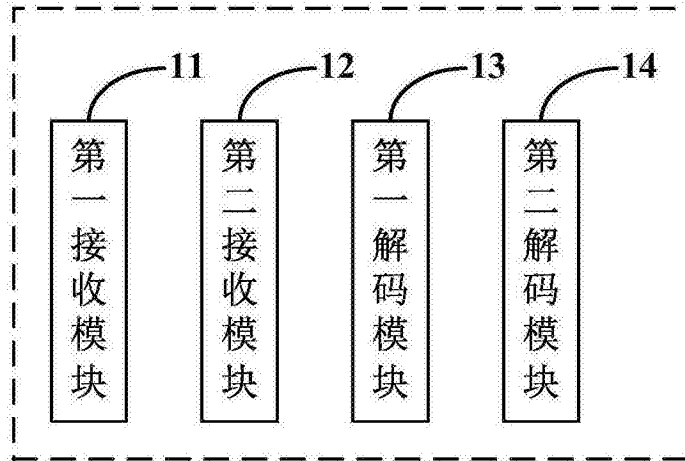


图2



图3



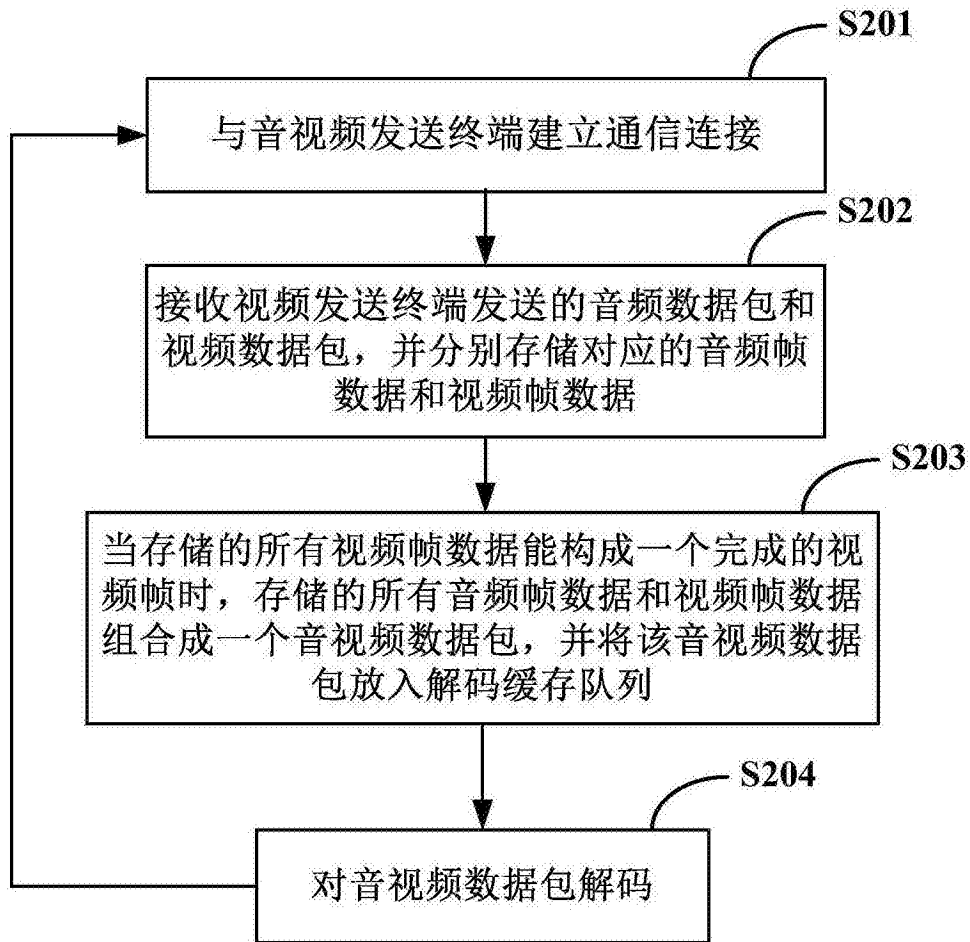


图4

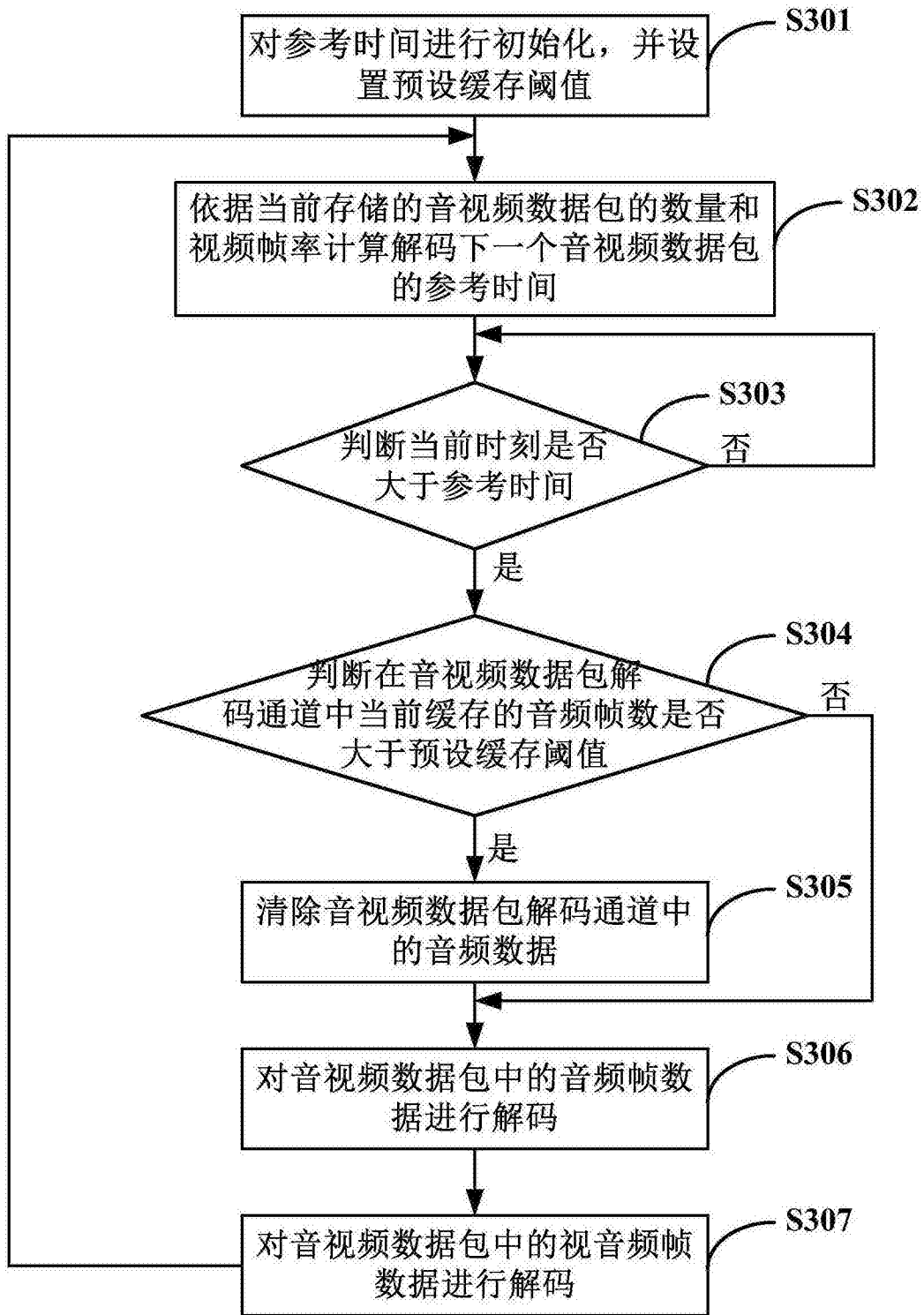


图5