



(12) 发明专利

(10) 授权公告号 CN 117896552 B

(45) 授权公告日 2024.07.12

(21) 申请号 202410290195.1

H04N 21/4402 (2011.01)

(22) 申请日 2024.03.14

H04N 7/15 (2006.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 117896552 A

(56) 对比文件

CN 113055742 A, 2021.06.29

CN 114996227 A, 2022.09.02

(43) 申请公布日 2024.04.16

审查员 席楠

(73) 专利权人 浙江华创视讯科技有限公司

地址 310051 浙江省杭州市滨江区滨安路  
1199号C9

(72) 发明人 吕少卿 俞鸣园 王克彦 曹亚曦  
费敏健

(74) 专利代理机构 深圳市威世博知识产权代理  
事务所(普通合伙) 44280

专利代理师 李秀云

(51) Int. Cl.

H04N 21/2343 (2011.01)

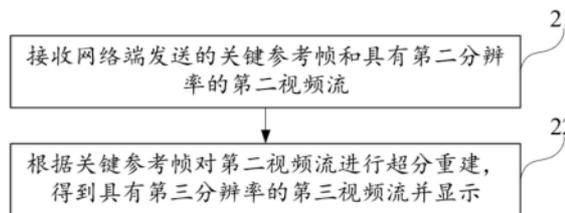
权利要求书2页 说明书13页 附图5页

(54) 发明名称

视频会议的处理方法、视频会议系统以及相关装置

(57) 摘要

本申请公开了视频会议的处理方法、视频会议系统以及相关装置。应用于接收端,该方法包括:接收网络端发送的关键参考帧和具有第二分辨率的第二视频流;其中,具有第二分辨率的第二视频流是网络端响应于网络质量低于预设质量,对发送端发送的具有第一分辨率的第一视频流进行分辨率降低得到;关键参考帧是网络端利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定得到;根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。通过上述方式,提升视频会议效果。



1. 一种视频会议的处理方法,其特征在于,应用于接收端,所述方法包括:

接收网络端发送的关键参考帧和具有第二分辨率的第二视频流;其中,所述具有第二分辨率的第二视频流是所述网络端响应于网络质量低于预设质量,对发送端发送的具有第一分辨率的第一视频流进行分辨率降低得到;所述关键参考帧是所述网络端利用所述第一视频流中每一视频帧的清晰度指数、细节保留度和内容重要性指数确定得到;

对于所述第二视频流中的每一视频帧,提取所述视频帧的第一关键特征点,以及提取所述关键参考帧中的第二关键特征点;匹配所述第一关键特征点和所述第二关键特征点,得到匹配特征点对;计算每一所述匹配特征点对之间的位置变化,得到运动信息;

基于所述运动信息,得到变形场;其中,所述变形场用于描述从每一视频帧至所述关键参考帧的特征变化;

利用所述变形场调整从所述关键参考帧提取的第一特征;

利用调整后的第一特征与每一视频帧的第二特征结合,进行超分重建,得到具有第三分辨率的第三视频流并显示。

2. 根据权利要求1所述的方法,其特征在于,显示所述具有第三分辨率的第三视频流之前,对所述具有第三分辨率的第三视频流进行质量评估。

3. 根据权利要求1所述的方法,其特征在于,所述利用调整后的第一特征与每一视频帧的第二特征结合,进行超分重建,包括:

利用生成对抗网络或卷积神经网络,结合调整后的第一特征与每一视频帧的第二特征进行超分重建。

4. 根据权利要求1所述的方法,其特征在于,所述基于所述运动信息,得到变形场,包括:

将所述运动信息应用于整个所述视频帧,得到所述视频帧对应的变形场。

5. 一种视频会议的处理方法,其特征在于,应用于网络端,所述方法包括:

接收发送端发送的具有第一分辨率的第一视频流,并监测网络质量;

响应于所述网络质量低于预设质量,降低所述第一视频流的分辨率,得到具有第二分辨率的第二视频流,以及利用所述第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出关键参考帧;

将所述第二视频流和所述关键参考帧发送至接收端,以使所述接收端对于所述第二视频流中的每一视频帧,提取所述视频帧的第一关键特征点,以及提取所述关键参考帧中的第二关键特征点;匹配所述第一关键特征点和所述第二关键特征点,得到匹配特征点对;计算每一所述匹配特征点对之间的位置变化,得到运动信息;基于所述运动信息,得到变形场;其中,所述变形场用于描述从每一视频帧至所述关键参考帧的特征变化;利用所述变形场调整从所述关键参考帧提取的第一特征;利用调整后的第一特征与每一视频帧的第二特征结合,进行超分重建,得到具有第三分辨率的第三视频流并显示。

6. 根据权利要求5所述的方法,其特征在于,所述响应于所述网络质量低于预设质量,降低所述第一视频流的分辨率,得到具有第二分辨率的第二视频流,包括:

响应于所述网络质量低于预设质量,结合所述关键参考帧确定下采样程度;所述下采样程度能够保证所述关键参考帧中的重要内容在超分重建过程中被识别;

基于所述下采样程度降低所述第一视频流的分辨率,得到具有第二分辨率的第二视频

流。

7. 根据权利要求5所述的方法,其特征在于,所述方法还包括:

接收所述接收端反馈信息,结合实时网络质量动态调整关键参考帧和超分辨率的选择策略。

8. 一种接收端,其特征在于,所述接收端包括存储器以及与所述存储器耦接的处理器和通信模组,所述存储器存储有至少一计算机程序,所述至少一计算机程序被所述处理器加载并执行时,用于实现如权利要求1-4任一项所述的方法。

9. 一种网络端,其特征在于,所述网络端包括存储器以及与所述存储器耦接的处理器和通信模组,所述存储器存储有至少一计算机程序,所述至少一计算机程序被所述处理器加载并执行时,用于实现如权利要求5-7任一项所述的方法。

10. 一种视频会议系统,其特征在于,所述视频会议系统包括:发送端、网络端和接收端,所述网络端如权利要求9所述的网络端,所述接收端如权利要求8所述的接收端。

11. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质有至少一段程序,所述至少一段程序被处理器加载并执行时,用于实现如权利要求1-7任一项所述的方法。

## 视频会议的处理方法、视频会议系统以及相关装置

### 技术领域

[0001] 本申请涉及视频会议技术领域,特别是涉及视频会议的处理方法、视频会议系统以及相关装置。

### 背景技术

[0002] 在视频会议中,动态场景(如快速移动的参与者或相机抖动)是常见的。相关技术在这些场景下往往难以保持视频帧间的连贯性和自然性,导致重建的视频出现模糊或不自然的现象。这种限制降低了视频会议的视觉质量和用户体验。

[0003] 以及视频会议通常依赖于网络连接,而网络带宽和稳定性可能因地理位置、网络拥堵或其他技术问题而波动。相关技术在网络条件不稳定时难以保持视频质量,尤其是在带宽有限的情况下,视频质量可能急剧下降。以及在视频会议中,保留关键视觉信息(如发言人的面部表情、重要的视觉提示)至关重要。然而,相关技术在重建过程中无法有效地保留这些关键信息,导致重要的非语言交流细节丢失。

### 发明内容

[0004] 本申请提供的视频会议的处理方法、视频会议系统以及相关装置,能够提升视频会议效果。

[0005] 为解决上述技术问题,本申请采用的一个技术方案是:提供一种视频会议的处理方法,应用于接收端,该方法包括:接收网络端发送的关键参考帧和具有第二分辨率的第二视频流;其中,具有第二分辨率的第二视频流是网络端响应于网络质量低于预设质量,对发送端发送的具有第一分辨率的第一视频流进行分辨率降低得到;关键参考帧是网络端利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定得到;根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0006] 其中,根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示,包括:对于第二视频流中的每一视频帧,估计每一视频帧与关键参考帧之间的运动信息;基于运动信息,得到变形场;其中,变形场用于描述从每一视频帧至关键参考帧的特征变化;利用变形场调整从关键参考帧提取的第一特征;利用调整后的第一特征与每一视频帧的第二特征结合,进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0007] 其中,显示具有第三分辨率的第三视频流之前,对具有第三分辨率的第三视频流进行质量评估。

[0008] 其中,利用调整后的第一特征与每一视频帧的第二特征结合,进行超分重建,包括:利用生成对抗网络或卷积神经网络,结合调整后的第一特征与每一视频帧的第二特征进行超分重建。

[0009] 其中,估计每一视频帧与关键参考帧之间的运动信息,包括:提取视频帧的第一关键特征点,以及提取关键参考帧中的第二关键特征点;匹配第一关键特征点和第二关键特征点,得到匹配特征点对;计算每一匹配特征点对之间的位置变化,得到运动信息。

[0010] 其中,基于运动信息,得到变形场,包括:将运动信息应用于整个视频帧,得到视频帧对应的变形场。

[0011] 为解决上述技术问题,本申请采用的另一技术方案是:提供一种视频会议的处理方法,应用于网络端,该方法包括:接收发送端发送的具有第一分辨率的第一视频流,并监测网络质量;响应于网络质量低于预设质量,降低第一视频流的分辨率,得到具有第二分辨率的第二视频流,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出关键参考帧;将第二视频流和关键参考帧发送至接收端,以使接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0012] 其中,响应于网络质量低于预设质量,降低第一视频流的分辨率,得到具有第二分辨率的第二视频流,包括:响应于网络质量低于预设质量,结合关键参考帧确定下采样程度;下采样程度能够保证关键参考帧中的重要内容在超分重建过程中被识别;基于下采样程度降低第一视频流的分辨率,得到具有第二分辨率的第二视频流。

[0013] 其中,该方法还包括:接收接收端反馈信息,结合实时网络质量动态调整关键参考帧和超分分辨率的选择策略。

[0014] 为解决上述技术问题,本申请采用的另一技术方案是:提供一种接收端,该接收端包括存储器以及与存储器耦接的处理器和通信模组,存储器存储有至少一计算机程序,至少一计算机程序被处理器加载并执行时,用于实现如应用于接收端的方法。

[0015] 为解决上述技术问题,本申请采用的另一技术方案是:提供一种网络端,该网络端包括存储器以及与存储器耦接的处理器和通信模组,存储器存储有至少一计算机程序,至少一计算机程序被处理器加载并执行时,用于实现如应用于网络端的方法。

[0016] 为解决上述技术问题,本申请采用的另一技术方案是:提供一种视频会议系统,该视频会议系统包括:发送端、网络端和接收端,网络端如上述的网络端,接收端如上述的接收端。

[0017] 为解决上述技术问题,本申请采用的另一技术方案是:提供一种计算机可读存储介质,该计算机可读存储介质有至少一段程序,至少一段程序被处理器加载并执行时,用于实现如上述任一技术方案提供的方法。

[0018] 区别于当前技术,本申请提供的视频会议的处理方法、视频会议系统以及相关装置,通过在视频传输过程中,当网络质量低于预设质量时,利用网络端对发送端发送的具有第一分辨率的第一视频流进行分辨率降低,得到具有第二分辨率的第二视频流,以此适应网络波动,保持视频流质量,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出第一视频流中的关键参考帧,然后在接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示,以此提升超分重建效果,进而提升视频会议效果。

## 附图说明

[0019] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。其中:

- [0020] 图1是本申请中视频会议系统一实施例的结构示意图；
- [0021] 图2是本申请中视频会议的处理方法一实施例的流程示意图；
- [0022] 图3是本申请中视频会议的处理方法另一实施例的流程示意图；
- [0023] 图4是本申请中步骤32一实施例的流程示意图；
- [0024] 图5是本申请中视频会议的处理方法另一实施例的流程示意图；
- [0025] 图6是本申请中视频会议的处理方法另一实施例的流程示意图；
- [0026] 图7是本申请中接收端一实施例的结构示意图；
- [0027] 图8是本申请中网络端一实施例的结构示意图；
- [0028] 图9是本申请中计算机可读存储介质一实施例的结构示意图。

### 具体实施方式

[0029] 下面结合附图和实施例,对本发明作进一步的详细描述。特别指出的是,以下实施例仅用于说明本发明,但不对本发明的范围进行限定。同样的,以下实施例仅为本发明的部分实施例而非全部实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0030] 在本文中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本发明的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是,本文所描述的实施例可以与其它实施例相结合。

[0031] 参阅图1,图1是本申请中视频会议系统一实施例的结构示意图。该视频会议系统100包括:发送端10、网络端20和接收端30。

[0032] 发送端10采集的视频流通过网络端20传输至接收端30进行显示,进而实现视频会议。

[0033] 可以理解,在实际应用过程中,因视频会议存在多方用户参与端中只有一个发送端,其余用户参与端均为接收端。在其他实施例中,视频会议中的每一参与端既是发送端也是接收端。

[0034] 在视频会议中,动态场景(如快速移动的参与者或相机抖动)是常见的。相关技术在这些场景下往往难以保持视频帧间的连贯性和自然性,导致重建的视频出现模糊或不自然的现象。这种限制降低了视频会议的视觉质量和用户体验。

[0035] 以及视频会议通常依赖于网络连接,而网络带宽和稳定性可能因地理位置、网络拥堵或其他技术问题而波动。相关技术在网络条件不稳定时难以保持视频质量,尤其是在带宽有限的情况下,视频质量可能急剧下降。以及在视频会议中,保留关键视觉信息(如发言人的面部表情、重要的视觉提示)至关重要。然而,相关技术在重建过程中无法有效地保留这些关键信息,导致重要的非语言交流细节丢失。

[0036] 视频会议可能涉及复杂的背景和多个参与者,这些场景的复杂性给传统超分辨率方法带来了挑战。这些方法可能无法有效地处理多人交互或频繁变化的背景,从而影响重建图像的质量。

[0037] 基于此,本申请提出通过在视频传输过程中,当网络质量低于预设质量时,利用网络端对发送端发送的具有第一分辨率的第一视频流进行分辨率降低,得到具有第二分辨率

的第二视频流,以此适应网络波动,保持视频流质量,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出第一视频流中的关键参考帧,然后在接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示,以此提升超分重建效果。具体参阅以下任一实施例。

[0038] 请参阅图2,图2是本申请中视频会议的处理方法一实施例的流程示意图。应用于接收端,该方法包括:

[0039] 步骤21:接收网络端发送的关键参考帧和具有第二分辨率的第二视频流。

[0040] 其中,具有第二分辨率的第二视频流是网络端响应于网络质量低于预设质量,对发送端发送的具有第一分辨率的第一视频流进行分辨率降低得到。即在视频流传输过程中,网络端实时监测网络质量,在网络质量低于预设质量时,对发送端发送的具有第一分辨率的第一视频流进行降分辨率处理得到。

[0041] 进一步,网络端利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定得到关键参考帧。

[0042] 步骤22:根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0043] 在一些实施例中,可以结合关键参考帧对具有第二分辨率的第二视频流进行超分重建,提升第二视频流的分辨率,得到具有第三分辨率的第三视频流并显示。

[0044] 在本实施例中,通过在视频传输过程中,当网络质量低于预设质量时,利用网络端对发送端发送的具有第一分辨率的第一视频流进行分辨率降低,得到具有第二分辨率的第二视频流,以此适应网络波动,保持视频流质量,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出第一视频流中的关键参考帧,然后在接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示,以此提升超分重建效果。

[0045] 请参阅图3,图3是本申请中视频会议的处理方法一实施例的流程示意图。应用于接收端,该方法包括:

[0046] 步骤31:接收网络端发送的关键参考帧和具有第二分辨率的第二视频流。

[0047] 其中,具有第二分辨率的第二视频流是网络端响应于网络质量低于预设质量,对发送端发送的具有第一分辨率的第一视频流进行分辨率降低得到;关键参考帧是网络端利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定得到。

[0048] 步骤32:对于第二视频流中的每一视频帧,估计每一视频帧与关键参考帧之间的运动信息。

[0049] 在一些实施例中,通过特征点匹配、光流算法或其他运动估计技术,计算出视频帧与关键参考帧之间的运动信息。

[0050] 在一些实施例中,参阅图4,步骤32可以是以下流程:

[0051] 步骤321:提取视频帧的第一关键特征点,以及提取关键参考帧中的第二关键特征点。

[0052] 在一些实施例中,可以利用相应的网络模型提取视频帧的第一关键特征点,以及提取关键参考帧中的第二关键特征点。

[0053] 步骤322:匹配第一关键特征点和第二关键特征点,得到匹配特征点对。

- [0054] 在一些实施例中,使用特征匹配算法,如FLANN(快速最近邻搜索库)或暴力匹配,以此匹配第一关键特征点和第二关键特征点,得到匹配特征点对。
- [0055] 在一些实施例中,可以计算第一关键特征点和第二关键特征点之间的相似度,选择最佳匹配。
- [0056] 可以理解,第一关键特征点的数量为多个。第二关键特征点的数量为多个。
- [0057] 步骤323:计算每一匹配特征点对之间的位置变化,得到运动信息。
- [0058] 在一些实施例中,可以使用光流算法或基于匹配的方法计算每一匹配特征点对中第一关键特征点和第二关键特征点之间的位置变化,得到运动信息。
- [0059] 步骤33:基于运动信息,得到变形场。
- [0060] 其中,变形场用于描述从每一视频帧至关键参考帧的特征变化。
- [0061] 在一些实施例中,可以将运动信息应用于整个视频帧,得到视频帧对应的变形场。
- [0062] 步骤34:利用变形场调整从关键参考帧提取的第一特征。
- [0063] 步骤35:利用调整后的第一特征与每一视频帧的第二特征结合,进行超分重建,得到具有第三分辨率的第三视频流并显示。
- [0064] 在一些实施例中,可以利用生成对抗网络或卷积神经网络,结合调整后的第一特征与每一视频帧的第二特征进行超分重建。在一些实施例中,超分重建的分辨率可以在网络端提前确定。
- [0065] 在一些实施例中,显示具有第三分辨率的第三视频流之前,对具有第三分辨率的第三视频流进行质量评估。在一些实施例中,使用图像质量评估工具和用户反馈来评估重建质量。进而可以根据质量评估结果调整超分辨率模型的参数,以此改善超分重建的效果。
- [0066] 在本实施例中,通过在视频传输过程中,当网络质量低于预设质量时,利用网络端对发送端发送的具有第一分辨率的第一视频流进行分辨率降低,得到具有第二分辨率的第二视频流,以此适应网络波动,保持视频流质量,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出第一视频流中的关键参考帧,然后在接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示,以此提升超分重建效果。
- [0067] 进一步,运动估计与变形场生成的结合为超分辨率重建提供了关键的运动信息。这种方法能够有效处理视频中的动态内容,如人物运动或相机移动。通过精确的运动估计和变形场生成,可以在重建高分辨率视频时保持更高的连贯性和自然性,尤其是在处理快速运动或复杂场景时。
- [0068] 请参阅图5,图5是本申请中视频会议的处理方法一实施例的流程示意图。应用于网络端,该方法包括:
- [0069] 步骤51:接收发送端发送的具有第一分辨率的第一视频流,并监测网络质量。
- [0070] 在视频流传输过程中,网络端实时监测网络质量,在网络质量低于预设质量时,对发送端发送的具有第一分辨率的第一视频流进行降分辨率处理得到。其中,可以通过监测延迟、丢包率和/或带宽,确定出网络质量。
- [0071] 步骤52:响应于网络质量低于预设质量,降低第一视频流的分辨率,得到具有第二分辨率的第二视频流。
- [0072] 步骤53:利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要

性指数确定出关键参考帧。

[0073] 关键参考帧通常是在网络质量良好时捕获的高分辨率帧。

[0074] 在一些实施例中,还可以结合网络质量和视频流的具体内容(如场景复杂度),动态选择相应的超分辨率,以便于在接收端按照此超分辨率进行超分重建。

[0075] 步骤54:将第二视频流和关键参考帧发送至接收端,以使接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0076] 在一些实施例中,可以利用压缩算法将第二视频流进行压缩发送至接收端,以减少数据传输量。接收端采用相应的解压缩算法对其进行解压,然后根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0077] 在本实施例中,通过在视频传输过程中,当网络质量低于预设质量时,利用网络端对发送端发送的具有第一分辨率的第一视频流进行分辨率降低,得到具有第二分辨率的第二视频流,以此适应网络波动,保持视频流质量,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出第一视频流中的关键参考帧,然后在接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示,以此提升超分重建效果。

[0078] 请参阅图6,图6是本申请中视频会议的处理方法一实施例的流程示意图。应用于网络端,该方法包括:

[0079] 步骤61:接收发送端发送的具有第一分辨率的第一视频流,并监测网络质量。

[0080] 步骤62:响应于网络质量低于预设质量,结合关键参考帧确定下采样程度。

[0081] 下采样程度能够保证关键参考帧中的重要内容在超分重建过程中被识别。

[0082] 步骤63:基于下采样程度降低第一视频流的分辨率,得到具有第二分辨率的第二视频流。

[0083] 步骤64:利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出关键参考帧。

[0084] 步骤65:将第二视频流和关键参考帧发送至接收端,以使接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0085] 在一些实施例中,设置反馈机制,接收接收端的反馈信息,结合实时网络质量动态调整关键参考帧和超分辨率的选择策略,以此更加适配当前的网络状态。

[0086] 在一应用场景中,本申请的视频会议系统的工作流程如下:

[0087] 在视频会议过程中,网络端进行网络质量监测与QoS触发。即在每个视频会议的视频流传输过程中,实时监测网络状况,包括延迟、丢包率和带宽。当网络质量下降到预设阈值时,触发质量服务(QoS)机制,启动超分功能。

[0088] 具体地,根据当前的网络状态(如带宽和延迟)和视频内容(如场景复杂度),动态选择合适的超分辨率。以及从视频流中选择关键参考帧。这通常是在网络状况良好时捕获的高分辨率帧,包含丰富的场景细节。

[0089] 具体地,上述超分辨率和关键参考帧的选择实现过程如下:

[0090] 首先进行网络状态监测。实时监测当前网络状态,包括带宽、延迟和丢包率。技术实现:使用网络监测工具,如Wireshark或自定义脚本,来捕获和分析网络流量和RTCP报告。

数据分析:计算网络的平均带宽( $B_{avg}$ ),延迟( $D_{avg}$ )和丢包率( $P_{loss}$ )。公式:例如, $B_{$

$avg = \Sigma(\text{带宽样本}) / \text{样本数量}$ 。这些参数将用于决定是否降低视频分辨率或触发超分重建。

[0091] 以及对视频内容分析。因场景复杂度高的视频可能需要更高的分辨率来保持细节,因此分析当前视频内容的复杂度。具体地,使用图像处理算法,如边缘检测(Canny算法)或纹理分析,来评估会议场景的复杂度。如,计算场景的纹理复杂度指数(T\_index)。对应的公式:例如, $T\_index = \Sigma(\text{像素纹理值}) / \text{总像素数}$ 。

[0092] 以及对画面内容分析。实时分析视频内容,识别重要人物或重要画面变换。如使用图像识别和机器学习技术来识别关键人物(如发言人或主持人)和重要事件。具体地,评估画面中人物的重要性和活动的显著性。以此可以根据画面内容的重要性动态调整超分分辨率和选择关键参考帧。

[0093] 以及场景变更检测。监测视频中的场景变化,如背景突然变更或画面剧烈切换。具体地,使用场景检测算法来识别画面中的显著变化。如计算场景变化的频率和幅度。通过上述方式,在场景发生重大变化时,重新选择关键参考帧以适应新场景。

[0094] 以及发言人和主持人变更检测。识别视频会议中的发言人和主持人变更。如使用语音识别和面部识别技术来识别当前的发言人或主持人。具体地,分析语音和面部数据,确定发言权的转移。如,在发言人或主持人变更时,调整超分分辨率或选择新的关键参考帧。

[0095] 以及接收端异常检测。在接收端检测视频流的异常情况,如画面冻结或质量突然下降。技术实现:使用视频质量评估工具来监测画面质量。数据分析:实时评估视频流的清晰度、流畅度和完整性。应用:在检测到异常时,请求新的关键参考帧或调整超分分辨率。

[0096] 以及动态调整策略。根据视频内容和接收端的反馈动态调整超分分辨率和关键参考帧。如,在网路端实现一个自适应系统,根据实时分析的结果调整超分策略。具体地,综合考虑网络状态、画面内容和接收端反馈动态调整超分分辨率和关键参考帧,以此确保视频会议中的视频质量最优化,特别是在关键时刻。

[0097] 反馈机制和网络适应:建立一个反馈机制,以便在网络状况变化时快速响应。技术实现:实时监测网络状况和接收端的反馈,快速调整超分策略。数据分析:持续评估网络带宽、延迟和丢包率。应用:在网络状况变化时,快速调整超分分辨率和关键参考帧的选择。

[0098] 上述的超分分辨率决策采用以下方式,基于网络状态、视频内容和画面重要性,决定超分分辨率。如使用决策树、机器学习模型或深度学习模型来确定最佳分辨率。模型应考虑网络参数、视频内容复杂度和画面中的关键元素(如重要人物)。具体地,综合考虑网络平均带宽(B\_avg)、平均延迟(D\_avg)、丢包率(P\_loss)、场景纹理复杂度指数(T\_index)和画面重要性指数(I\_index)。可以采用以下公式:分辨率 = f(B\_avg, D\_avg, P\_loss, T\_index, I\_index)。即选择一种分辨率,既能适应网络条件,又能保持视频内容的清晰度,并突出画面中的重要元素。

[0099] 上述的关键参考帧的选择采用以下方式:基于网络状况、视频内容和画面动态变化选择高分辨率的关键参考帧。具体地,分析视频流中的每一帧,考虑清晰度、细节保留度和画面内容的重要性(如发言人变更、场景变化)。如评估每帧的清晰度指数(C\_index)、细节保留度(D\_score)和内容重要性指数(CI\_index)。如,参考帧得分 = g(C\_index, D\_score, CI\_index)。其中,可以选择得分最高的帧作为关键参考帧,以确保在超分重建过程中保留关键内容和高质量的细节。

[0100] 在网络质量不佳时,将视频流降级为低分辨率版本以适应带宽限制。低分辨率视频帧通过压缩算法传输,以减少所需的数据量。

[0101] 具体地,低分辨率视频帧的处理实现过程如下:

[0102] 关键参考帧的利用:在确定进行低分辨率处理时,首先考虑已选择的关键参考帧。具体地,使用图像处理算法比较当前帧与关键参考帧,识别重要的视觉元素。如分析两帧之间的差异,特别是在重要元素(如人脸、标志性物体)上的差异,以此确保在下采样过程中,这些重要元素的细节尽可能保留。

[0103] 动态下采样决策:根据当前网络状况和关键参考帧的内容,动态决定下采样的程度。具体地,结合网络带宽、延迟、丢包率和关键参考帧的内容特征来做出决策。如考虑网络参数和关键参考帧的视觉重要性,以此在保证重要内容可识别的前提下,选择合适的下采样比例。

[0104] 执行下采样:对当前视频帧执行下采样处理。具体地,使用高级图像处理算法(如双三次插值)进行下采样,同时考虑保留关键元素的细节,以此在生成低分辨率版本的视频帧,同时尽量保留关键元素的清晰度。

[0105] 低分辨率帧的质量增强:对下采样后的视频帧进行质量增强处理。具体地,使用图像增强技术,如锐化、对比度调整,来提升低分辨率帧的视觉质量,以此改善低分辨率视频帧的视觉效果,使其在较低的分辨率下仍然清晰可见。

[0106] 压缩与编码:对处理后的低分辨率视频帧进行压缩和编码。技术实现:使用高效的视频编码技术(如H.265)进行压缩,以减少数据量,以此确保视频帧在网络上传输时占用的带宽最小。

[0107] 接收端接收到低分辨率视频流后,根据预先选择的关键参考帧进行超分重建。如果关键参考帧与当前帧存在显著差异,或因网络问题导致的帧损坏,接收端可能需要采取补偿措施,如请求新的参考帧。

[0108] 接收端对于每个低分辨率帧,估计其与高分辨率参考帧(关键参考帧)之间的运动。基于估计的运动信息,生成变形场,描述从参考帧到目标帧的特征变化。

[0109] 接收端使用变形场调整从高分辨率参考帧提取的特征,以匹配低分辨率目标帧。将调整后的特征与低分辨率帧的特征结合,使用神经网络进行超分辨率重建。

[0110] 接收端在超分重建后,评估重建视频帧的质量。如果质量不佳或存在明显的重建错误,进行进一步的调整或重新请求参考帧。考虑到实时性要求,这一步骤需要快速高效地执行。

[0111] 接收端将重建的高分辨率视频流输出到会议室的显示设备上,以提供更清晰的视频会议体验。

[0112] 其中,接收端超分重建的实现过程如下:

[0113] 接收低分辨率视频帧:接收端首先接收通过网络传输的低分辨率视频帧。具体地,使用网络协议(如RTP/RTCP)接收并解码视频流。如将接收到的压缩数据解码为可用于处理的视频帧格式,以此确保视频帧在接收后能够被正确解析和处理。

[0114] 帧质量评估:评估接收到的低分辨率视频帧的质量。具体地,使用图像质量评估算法(如PSNR或SSIM)来评估视频帧的质量。如,计算每帧的质量得分。如采用以下公式,  $PSNR = 10 * \log_{10}(MAX\_I^2 / MSE)$ 。以此确定哪些帧需要进行超分辨率处理。

[0115] 关键参考帧的匹配:将接收到的低分辨率帧与选定的关键参考帧进行匹配。具体地,使用图像匹配算法(如特征匹配或光流估计)来找到关键参考帧与当前帧之间的对应关系。如计算参考帧和当前帧之间的相似度或运动向量。例如,采用光流估计公式或特征匹配算法。关键参考帧的匹配能够确保超分重建过程中参考帧与当前帧的一致性。

[0116] 运动估计与变形场生成:基于参考帧和当前帧之间的匹配,估计运动并生成变形场。具体地,使用光流算法或其他运动估计技术来估计帧间运动。如生成描述帧间运动的变形场。例如,采用光流估计公式或变形场计算公式,以此为超分重建过程提供必要的运动和变形信息。

[0117] 超分辨率重建:使用深度学习模型或其他超分辨率技术对低分辨率帧进行重建。具体地,应用如生成对抗网络(GAN)或卷积神经网络(CNN)的超分辨率模型。如结合低分辨率帧、变形场和关键参考帧的信息进行重建。通过超分重建生成接近高分辨率的视频帧,提升视觉质量。

[0118] 重建质量评估与优化:评估重建后的视频帧质量,并进行必要的优化。具体地,使用图像质量评估工具和用户反馈来评估重建质量。如根据质量评估结果调整超分辨率模型的参数,以此确保重建的视频帧在视觉上满足用户需求。

[0119] 上述运动估计与变形场生成的实现过程如下:

[0120] 运动估计的准备:准备进行运动估计的初始工作,包括收集必要的数据和设置参数。具体地,从接收到的视频流中提取当前低分辨率帧和关键参考帧。如确保两帧的格式和尺寸一致,以便于进行运动估计。以此为运动估计创建一个稳定的起始环境,确保后续步骤的准确性。

[0121] 特征提取:从当前帧和关键参考帧中提取特征点。具体地,使用特征提取算法,如SIFT(尺度不变特征变换)或ORB(Oriented FAST and Rotated BRIEF)。如识别两帧中的关键特征点,这些点应在两帧中都清晰可见。例如,采用SIFT算法中的特征点检测和描述子生成公式。以此提取的特征点将用于后续的匹配和运动估计。

[0122] 特征点匹配:将当前帧中的特征点与关键参考帧中的特征点进行匹配。具体地,使用特征匹配算法,如FLANN(快速最近邻搜索库)或暴力匹配。如,计算两帧中特征点之间的相似度,选择最佳匹配。例如,采用基于描述子的特征点匹配公式,以此确保正确匹配特征点,为准确估计运动提供基础。

[0123] 运动向量计算:基于匹配的特征点计算运动向量。具体地,使用光流算法或基于匹配的方法计算每对匹配点之间的运动。如计算每对匹配点之间的位置变化,生成运动向量。例如,采用光流估计中的Lucas-Kanade方法或Horn-Schunck算法。其中,运动向量描述了参考帧到当前帧的运动,是生成变形场的关键。

[0124] 变形场的构建:使用运动向量构建变形场。具体地,将运动向量应用于整个帧,生成一个描述整个帧变形的场。如,分析运动向量的分布和强度,构建变形场。例如,采用基于运动向量的变形场生成公式。其中,变形场将用于指导超分辨率重建过程中的像素变换。

[0125] 变形场的优化和应用:优化变形场并将其应用于超分辨率重建过程。具体地,使用插值和平滑技术优化变形场,减少不连续和异常。如确保变形场在整个帧上平滑且连续。应用优化后的变形场于超分辨率重建,以实现更自然和连贯的视觉效果。

[0126] 上述的特征融合与重建的实现过程如下:

[0127] 特征提取:从低分辨率视频帧和关键参考帧中提取特征。具体地,使用深度学习模型,如卷积神经网络(CNN),提取两帧的特征。如将视频帧输入到CNN中,提取特征图(feature maps)。例如,采用特征提取公式  $F(x) = \text{ReLU}(W * x + b)$ ,其中  $W$  和  $b$  是卷积层的权重和偏置。其中,提取的特征将用于后续的融合和重建步骤。

[0128] 特征融合:将低分辨率帧的特征与变形场调整后的关键参考帧特征融合。具体地,使用特征融合技术,如特征叠加或特征加权融合。如结合两组特征,创建一个综合特征表示。例如,采用融合公式  $G(f1, f2) = \alpha * f1 + (1 - \alpha) * f2$ ,其中  $f1$  和  $f2$  是两组特征, $\alpha$  是融合系数。其中,融合的特征将提供更丰富的信息,用于重建高质量的视频帧。

[0129] 运动补偿:对融合的特征进行运动补偿,以考虑帧间的运动。具体地,应用变形场于融合的特征,以补偿运动。如根据变形场调整特征图的位置。例如,采用运动补偿公式  $M(f, v) = f(x + v)$ ,其中  $f$  是特征图, $v$  是变形场中的运动向量。其中,运动补偿确保特征在时间上的连续性和一致性。

[0130] 超分辨率重建:使用融合和运动补偿后的特征进行超分辨率重建。具体地,应用深度学习模型,如生成对抗网络(GAN),进行重建。如将融合的特征输入到GAN的生成器中,生成高分辨率帧。例如,采用GAN生成器公式  $G(f) = \tanh(Wg * f + bg)$ ,其中  $Wg$  和  $bg$  是生成器的权重和偏置。其中,生成的高分辨率帧应接近于原始的高分辨率质量。

[0131] 重建质量评估:评估重建后的视频帧质量。具体地,使用图像质量评估工具,如PSNR或SSIM,评估重建帧的质量。如,计算重建帧与原始高分辨率帧之间的质量差异。公式:例如,SSIM公式  $SSIM(x, y) = (2\mu_x\mu_y + C1) (2\sigma_{xy} + C2) / (\mu_x^2 + \mu_y^2 + C1) (\sigma_x^2 + \sigma_y^2 + C2)$ 。其中,重建质量评估能够确保重建的视频帧在视觉上满足质量标准。

[0132] 反馈和优化:根据重建质量的评估结果进行反馈和优化。具体地,调整超分辨率模型的参数或重新训练模型以改善重建质量。如分析质量评估结果,识别重建过程中的问题。其中,持续优化超分辨率重建过程,以适应不同的视频内容和网络条件。

[0133] 具体地,运动估计与变形场生成过程如下:

[0134] 运动估计:这一步骤涉及分析低分辨率视频帧与高分辨率关键参考帧之间的运动。通过特征点匹配、光流算法或其他运动估计技术,计算出帧间的运动向量。

[0135] 变形场生成:基于运动估计的结果,生成一个变形场,这个变形场描述了从关键参考帧到当前帧的像素级运动。变形场是一个向量场,每个向量指示特定像素或特征点在两帧之间的移动。

[0136] 具体地,特征融合与重建过程如下:

[0137] 特征提取:从低分辨率帧和高分辨率关键参考帧中提取特征。这通常通过深度学习模型(如CNN)完成,以获取两帧的特征表示。

[0138] 应用变形场:将变形场应用于关键参考帧的特征。这意味着根据变形场调整参考帧的特征,使其与当前帧的几何和运动特性对齐。

[0139] 特征融合:将调整后的关键参考帧特征与当前低分辨率帧的特征结合。这个融合过程考虑了两帧的内容和运动信息,创建了一个综合的特征表示。

[0140] 重建高分辨率帧:使用融合的特征通过超分辨率模型(如GAN或CNN)重建出高分辨率的视频帧。这个过程利用了融合特征中的丰富信息来生成清晰、细节丰富的高分辨率图像。

[0141] 结合过程如下:

[0142] 运动一致性:通过运动估计与变形场生成,确保关键参考帧的特征在几何和运动上与当前帧一致。这对于保持视频帧间的自然过渡和连贯性至关重要。

[0143] 信息增强:特征融合过程结合了调整后的参考帧特征和当前帧特征,增强了重建过程中的信息量。这使得重建的高分辨率帧能够更好地恢复丢失的细节和纹理。

[0144] 质量提升:最终,这种结合运动信息和特征信息的方法能够显著提升重建帧的质量,特别是在动态场景和复杂运动情况下。

[0145] 在本应用场景中,在超分辨率重建中,智能选择关键参考帧,这不仅基于图像质量,还考虑了画面内容的重要性和场景的动态变化。通过选择最合适的关键参考帧,可以确保在重建过程中保留关键视觉信息,特别是在动态场景和快速内容变化的情况下。

[0146] 以及根据实时网络状况和视频内容动态调整超分辨率重建策略。这种方法使得超分重建过程更加灵活,能够适应网络带宽的波动和视频内容的变化。这种动态调整机制提高了超分辨率重建的适应性和效率,特别是在网络条件不稳定的远程视频会议场景中。

[0147] 进一步,运动估计与变形场生成的结合为超分辨率重建提供了关键的运动信息。这种方法能够有效处理视频中的动态内容,如人物运动或相机移动。通过精确的运动估计和变形场生成,可以在重建高分辨率视频时保持更高的连贯性和自然性,尤其是在处理快速运动或复杂场景时。

[0148] 这些效果共同作用于提升低分辨率视频帧的质量,特别适用于网络条件变化大的远程视频会议场景。它们不仅提高了视频质量,还增强了模型对于视频内容变化的适应性,使得超分辨率重建过程更加高效和准确。

[0149] 参阅图7,图7是本申请中接收端一实施例的结构示意图。该接收端30包括处理器31以及与处理器31耦接的存储器32和通信模组33,存储器32存储有至少一计算机程序,至少一计算机程序被处理器加载并执行时,用于实现以下方法:

[0150] 接收网络端发送的关键参考帧和具有第二分辨率的第二视频流;其中,具有第二分辨率的第二视频流是网络端响应于网络质量低于预设质量,对发送端发送的具有第一分辨率的第一视频流进行分辨率降低得到;关键参考帧是网络端利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定得到;根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0151] 可以理解,至少一计算机程序被处理器31加载并执行时,还用于实现上述任一实施例的方法。

[0152] 参阅图8,图8是本申请中网络端一实施例的结构示意图。该网络端20包括处理器21以及与处理器21耦接的存储器22和通信模组23,存储器22存储有至少一计算机程序,至少一计算机程序被处理器加载并执行时,用于实现以下方法:

[0153] 接收发送端发送的具有第一分辨率的第一视频流,并监测网络质量;响应于网络质量低于预设质量,降低第一视频流的分辨率,得到具有第二分辨率的第二视频流,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出关键参考帧;将第二视频流和关键参考帧发送至接收端,以使接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0154] 可以理解,至少一计算机程序被处理器21加载并执行时,还用于实现上述任一实

施例的方法。

[0155] 请参阅图9,图9是本申请中计算机可读存储介质一实施例的结构示意图。该计算机可读存储介质110有至少一段程序111,至少一段程序111被处理器加载并执行时,用于实现以下方法:

[0156] 接收网络端发送的关键参考帧和具有第二分辨率的第二视频流;其中,具有第二分辨率的第二视频流是网络端响应于网络质量低于预设质量,对发送端发送的具有第一分辨率的第一视频流进行分辨率降低得到;关键参考帧是网络端利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定得到;根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示;

[0157] 或者,接收发送端发送的具有第一分辨率的第一视频流,并监测网络质量;响应于网络质量低于预设质量,降低第一视频流的分辨率,得到具有第二分辨率的第二视频流,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出关键参考帧;将第二视频流和关键参考帧发送至接收端,以使接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示。

[0158] 可以理解,至少一段程序111被处理器加载并执行时,用于实现上述任一实施例的方法。

[0159] 综上所述,本申请提供的视频会议的处理方法、视频会议系统以及相关装置,通过在视频传输过程中,当网络质量低于预设质量时,利用网络端对发送端发送的具有第一分辨率的第一视频流进行分辨率降低,得到具有第二分辨率的第二视频流,以此适应网络波动,保持视频流质量,以及利用第一视频流中每一视频帧的清晰度指数、细节保留度和/或内容重要性指数确定出第一视频流中的关键参考帧,然后在接收端根据关键参考帧对第二视频流进行超分重建,得到具有第三分辨率的第三视频流并显示,以此提升超分重建效果,进而提升视频会议效果。

[0160] 在本申请所提供的几个实施方式中,应该理解到,所揭露的系统,装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施方式仅仅是示意性的,例如,所述模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0161] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施方式方案的目的。

[0162] 另外,在本发明各个实施方式中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0163] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式

体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)或处理器(processor)执行本发明各个实施方式所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0164] 以上所述仅为本发明的实施方式,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

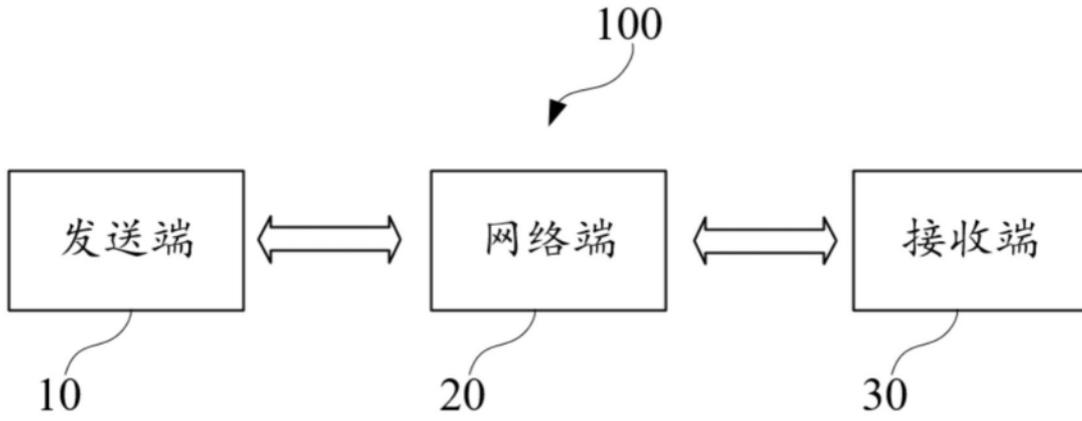


图1

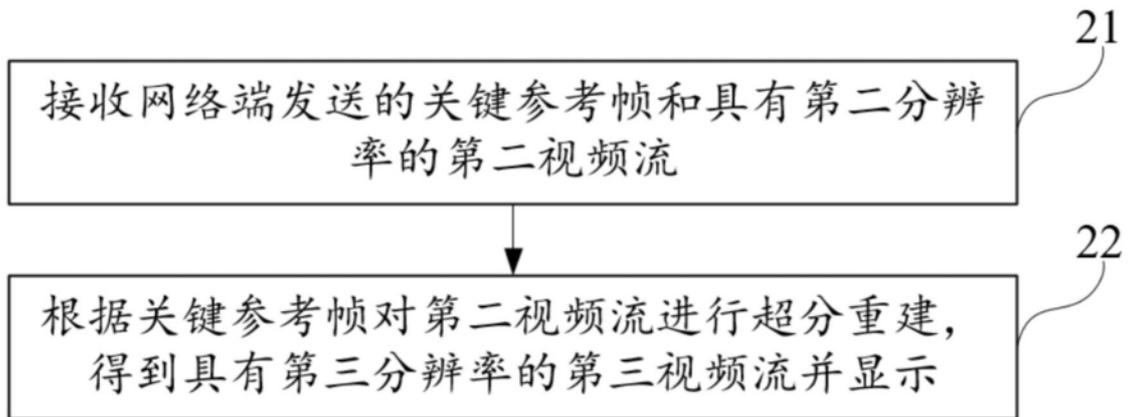


图2

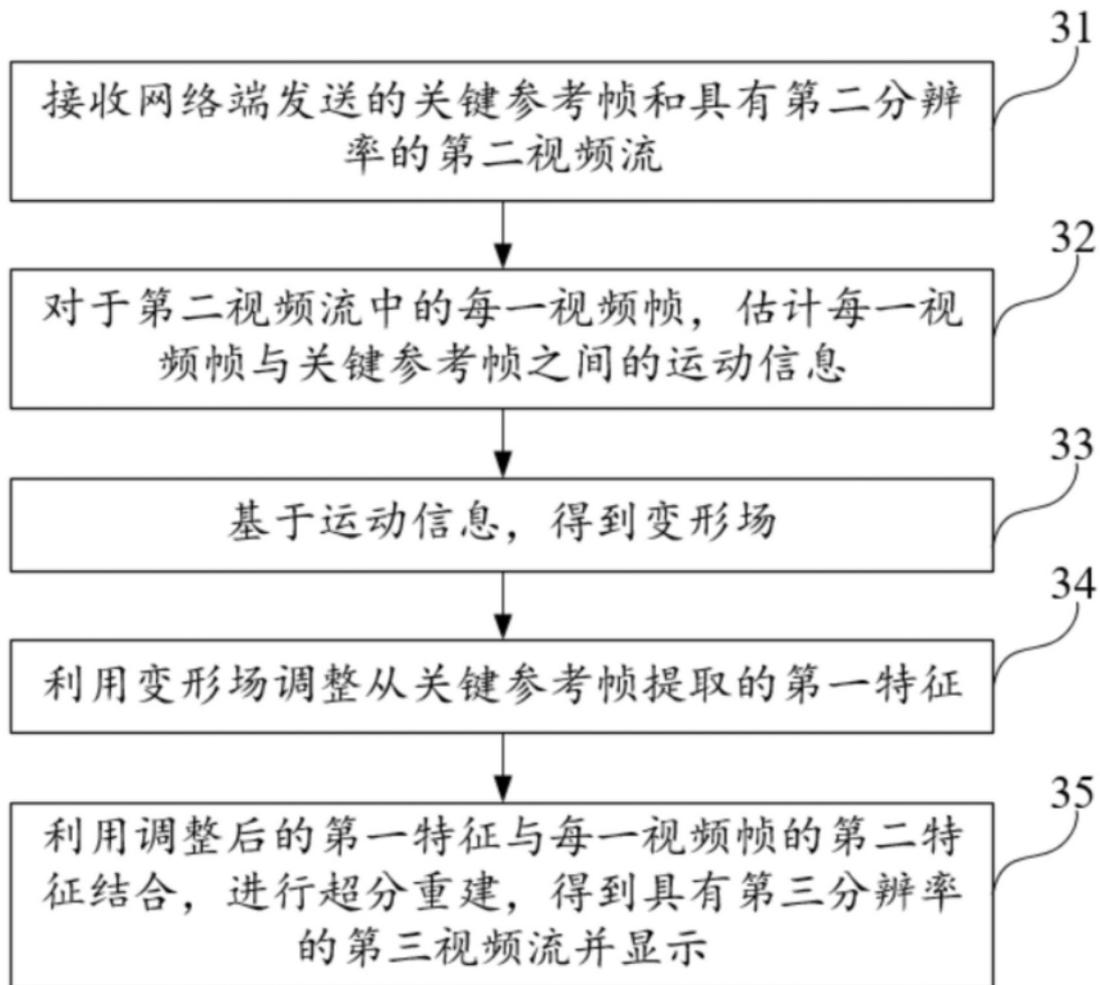


图3

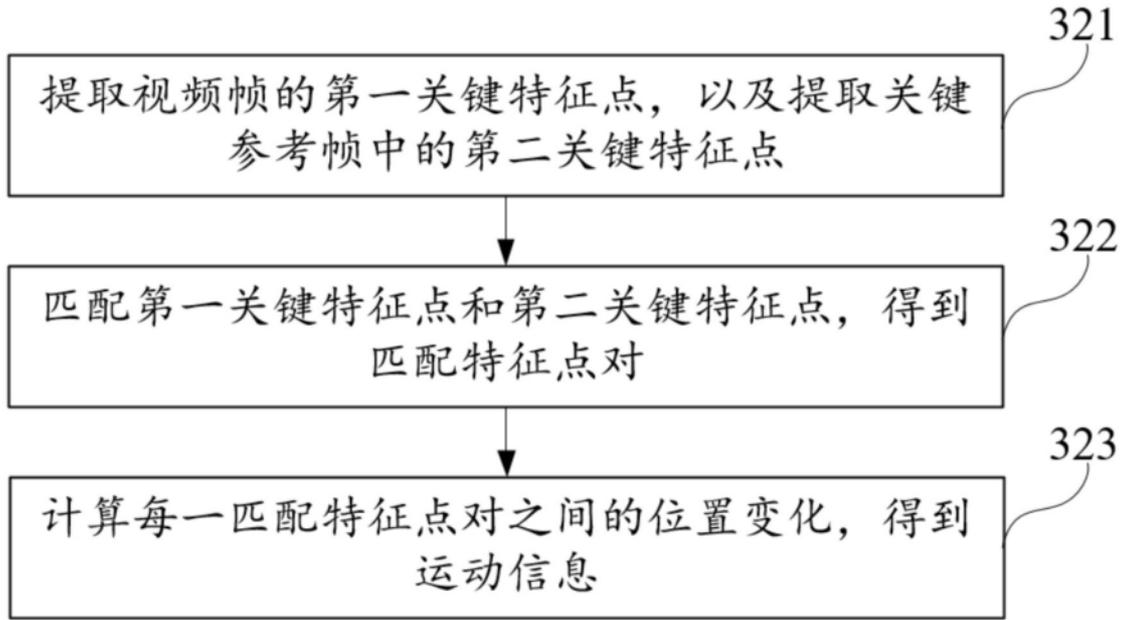


图4

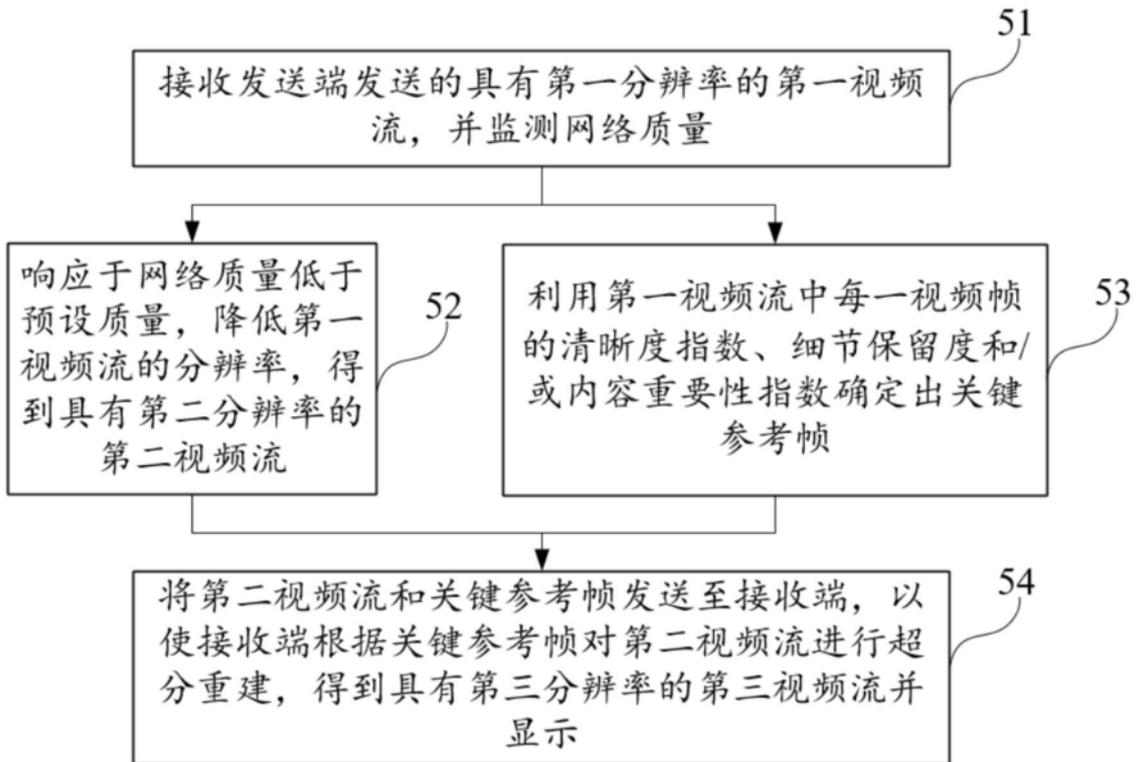


图5

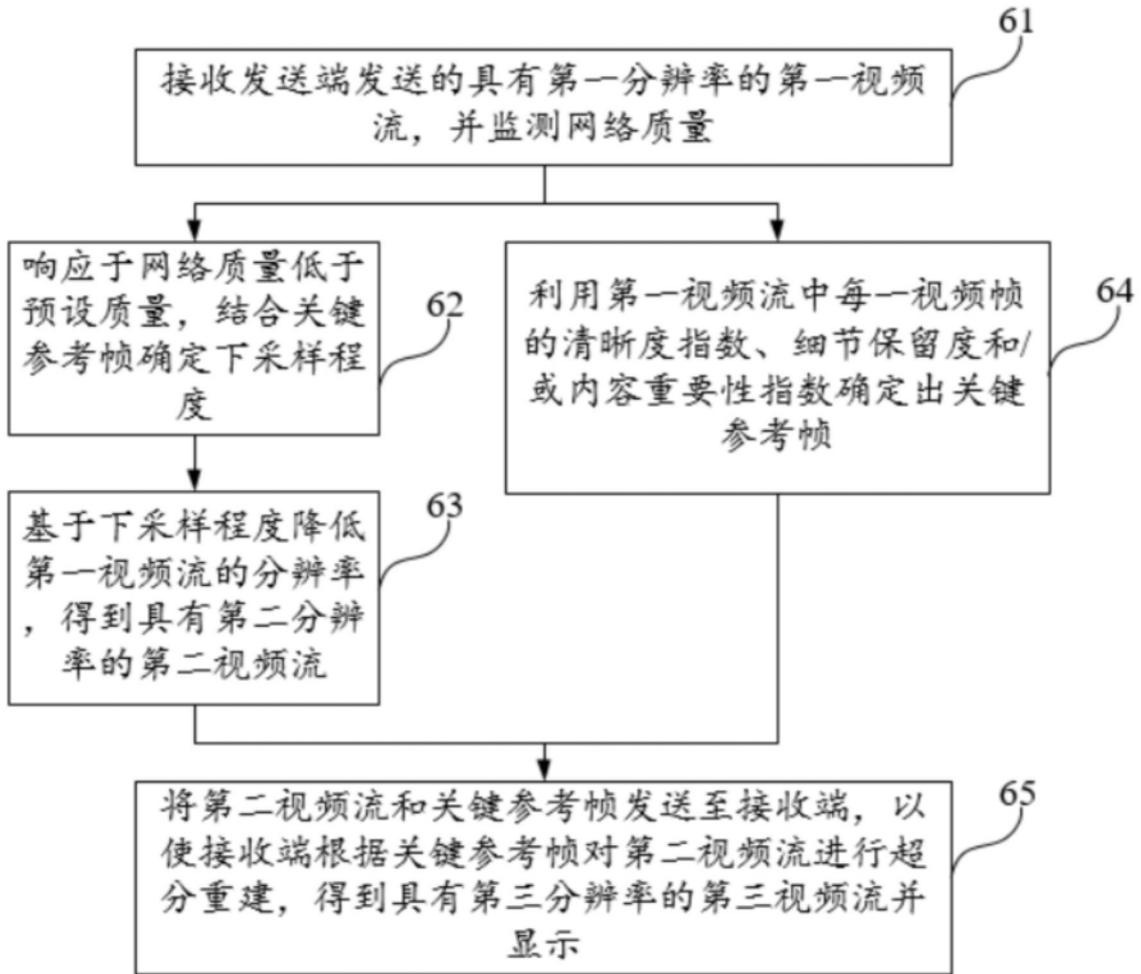


图6

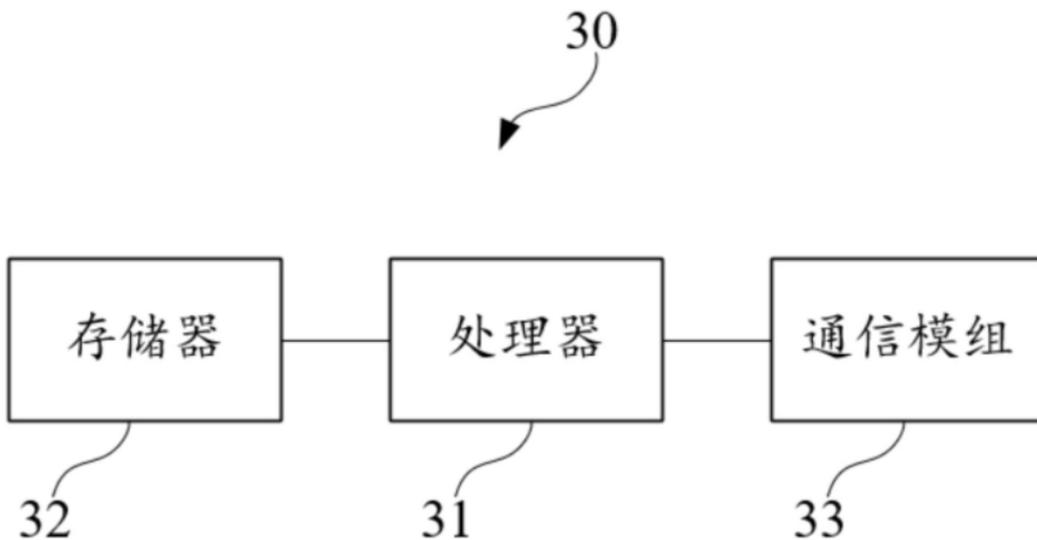


图7

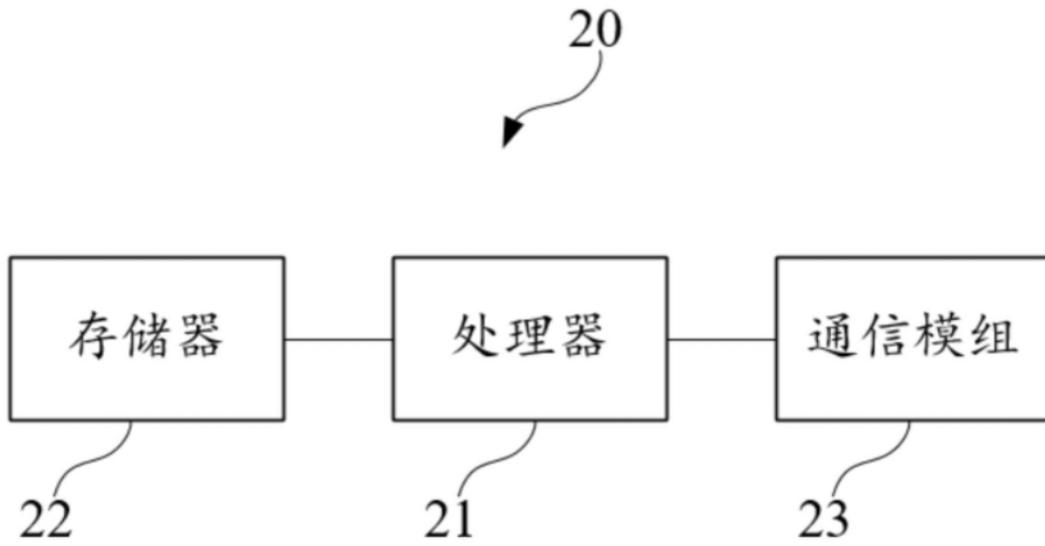


图8

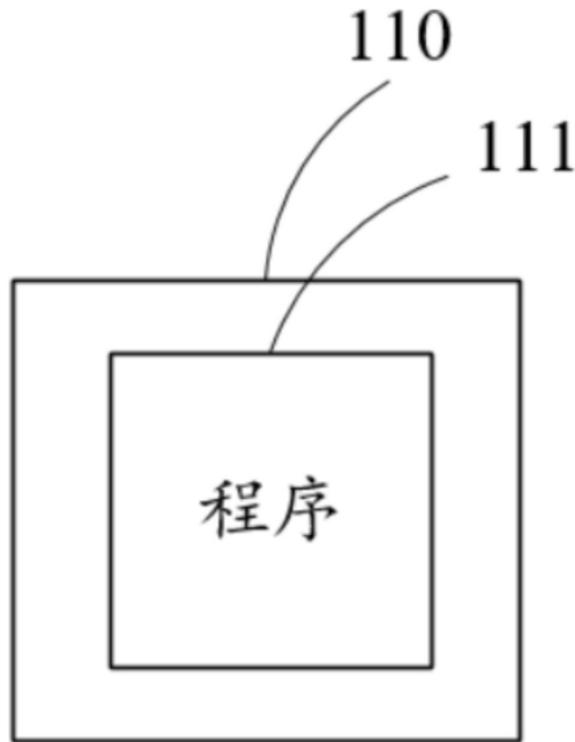


图9