



(12) 发明专利申请

(10) 申请公布号 CN 104967891 A

(43) 申请公布日 2015. 10. 07

(21) 申请号 201510370113. 5

H04N 21/439(2011. 01)

(22) 申请日 2015. 06. 29

(71) 申请人 高翔

地址 100013 北京市东城区和平里五区甲 4
楼 2 门 502 号

(72) 发明人 高翔 谢灿豪

(74) 专利代理机构 北京同立钧成知识产权代理
有限公司 11205

代理人 娄冬梅 黄健

(51) Int. Cl.

H04N 21/4147(2011. 01)

H04N 21/43(2011. 01)

H04N 5/92(2006. 01)

H04N 5/76(2006. 01)

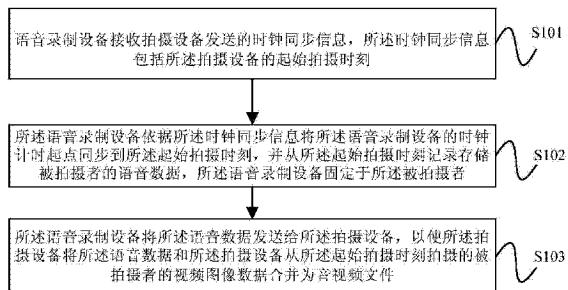
权利要求书4页 说明书10页 附图3页

(54) 发明名称

音视频文件生成方法及装置

(57) 摘要

本发明实施例提供一种音视频文件生成方法及装置。该方法包括：语音录制设备接收拍摄设备发送的时钟同步信息；语音录制设备依据时钟同步信息将语音录制设备的时钟计时起点同步到起始拍摄时刻，并从起始拍摄时刻记录存储被拍摄者的语音数据；拍摄设备将语音数据和视频图像数据合并为音视频文件。本发明实施例通过语音数据和视频图像数据对应相同的起始拍摄时刻即基于相同的时间轴，即使被拍摄者处于运动状态且远离拍摄者时，固定于被拍摄者的语音录制设备能够记录存储清晰的语音数据，保证语音数据和视频图像数据具有很高的匹配度，并保证合并后的音视频文件具有很高的语音质量。



1. 一种音视频文件生成方法,其特征在于,包括:

语音录制设备接收拍摄设备发送的时钟同步信息,所述时钟同步信息包括所述拍摄设备的起始拍摄时刻;

所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻,并从所述起始拍摄时刻记录存储被拍摄者的语音数据,所述语音录制设备固定于所述被拍摄者;

所述语音录制设备将所述语音数据发送给所述拍摄设备,以使所述拍摄设备将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

2. 根据权利要求 1 所述的方法,其特征在于,所述语音录制设备接收拍摄设备发送的时钟同步信息之前,还包括:

所述语音录制设备接收所述拍摄设备发送的初始时钟信息,所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳;

所述语音录制设备依据所述时间戳开始计时。

3. 根据权利要求 2 所述的方法,其特征在于,所述语音录制设备依据所述时间戳开始计时之后,还包括:

所述语音录制设备向所述拍摄设备发送拍摄启动请求,以使所述拍摄设备启动拍摄功能;

所述语音录制设备将所述语音数据发送给所述拍摄设备之后,还包括:

所述语音录制设备向所述拍摄设备发送拍摄结束请求,以使所述拍摄设备关闭拍摄功能。

4. 根据权利要求 1-3 任一项所述的方法,其特征在于,所述语音录制设备接收拍摄设备发送的时钟同步信息,包括:

所述语音录制设备通过无线通信的方式接收拍摄设备发送的时钟同步信息;

所述语音录制设备将所述语音数据发送给所述拍摄设备,包括:

所述语音录制设备将所述语音数据通过无线通信的方式发送给所述拍摄设备。

5. 根据权利要求 4 所述的方法,其特征在于,还包括:

所述语音录制设备监测其自身的开关状态、电量状态和存储空间状态,并生成状态信息,所述状态信息至少包括:开关状态信息、电量状态信息和存储空间状态信息;

所述语音录制设备将所述状态信息通过无线通信方式发送给所述拍摄设备,以使所述拍摄设备依据所述状态信息控制所述语音录制设备开启或关闭。

6. 根据权利要求 5 所述的方法,其特征在于,所述从所述起始拍摄时刻记录存储被拍摄者的语音数据之后,还包括:

所述语音录制设备对音量大于预设音量的语音数据进行 AGC 限幅处理;

所述语音录制设备的初始时钟与 GPS 时钟保持同步,所述拍摄设备的初始时钟与 GPS 时钟或所述拍摄设备所属基站的时钟保持同步。

7. 一种音视频文件生成方法,其特征在于,包括:

拍摄设备向语音录制设备发送时钟同步信息,所述时钟同步信息包括所述拍摄设备的起始拍摄时刻,以使所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时

钟时起点同步到所述起始拍摄时刻，并从所述起始拍摄时刻记录存储被拍摄者的语音数据，所述语音录制设备固定于所述被拍摄者；

所述拍摄设备接收所述语音录制设备发送的所述语音数据，并将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

8. 根据权利要求 7 所述的方法，其特征在于，所述拍摄设备向语音录制设备发送时钟同步信息之前，还包括：

所述拍摄设备向语音录制设备发送初始时钟信息，所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳，以使所述语音录制设备依据所述时间戳开始计时。

9. 根据权利要求 8 所述的方法，其特征在于，所述语音数据和所述视频图像数据具有相同的时间轴；

所述将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件，包括：

依据所述相同的时间轴将所述语音数据和所述视频图像数据合并为音视频文件。

10. 根据权利要求 9 所述的方法，其特征在于，所述拍摄设备向语音录制设备发送时钟同步信息之前，还包括：

所述拍摄设备接收所述语音录制设备发送的拍摄启动请求，并依据所述拍摄启动请求启动拍摄功能；

所述拍摄设备接收所述语音录制设备发送的所述语音数据之后，还包括：

所述拍摄设备接收所述语音录制设备发送的拍摄结束请求，并依据所述拍摄结束请求关闭拍摄功能。

11. 根据权利要求 10 所述的方法，其特征在于，所述拍摄设备向语音录制设备发送时钟同步信息，包括：

所述拍摄设备通过无线通信的方式向语音录制设备发送时钟同步信息；

所述拍摄设备接收所述语音录制设备发送的所述语音数据，包括：

所述拍摄设备通过无线通信的方式接收所述语音录制设备发送的所述语音数据。

12. 根据权利要求 11 所述的方法，其特征在于，还包括：

所述拍摄设备通过无线通信方式接收所述语音录制设备发送的状态信息，并依据所述状态信息控制所述语音录制设备开启或关闭，所述状态信息至少包括：开关状态信息、电量状态信息和存储空间状态信息。

13. 一种语音录制设备，其特征在于，包括：

接收模块，用于接收拍摄设备发送的时钟同步信息，所述时钟同步信息包括所述拍摄设备的起始拍摄时刻；

同步模块，用于设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻；

记录存储模块，用于从所述起始拍摄时刻记录存储被拍摄者的语音数据，所述语音录制设备固定于所述被拍摄者；

发送模块，用于将所述语音数据发送给所述拍摄设备，以使所述拍摄设备将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频

文件。

14. 根据权利要求 13 所述的语音录制设备，其特征在于，所述接收模块还用于接收所述拍摄设备发送的初始时钟信息，所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳；

所述语音录制设备还包括计时模块，所述计时模块用于依据所述时间戳开始计时。

15. 根据权利要求 14 所述的语音录制设备，其特征在于，所述发送模块还用于所述语音录制设备依据所述时间戳开始计时之后，向所述拍摄设备发送拍摄启动请求，以使所述拍摄设备启动拍摄功能；所述语音录制设备将所述语音数据发送给所述拍摄设备之后，向所述拍摄设备发送拍摄结束请求，以使所述拍摄设备关闭拍摄功能。

16. 根据权利要求 13-15 任一项所述的语音录制设备，其特征在于，所述接收模块具体用于通过无线通信的方式接收拍摄设备发送的时钟同步信息；

所述发送模块具体用于将所述语音数据通过无线通信的方式发送给所述拍摄设备。

17. 根据权利要求 16 所述的语音录制设备，其特征在于，还包括：

监测模块，用于监测其自身的开关状态、电量状态和存储空间状态，并生成状态信息，所述状态信息至少包括：开关状态信息、电量状态信息和存储空间状态信息；

所述发送模块还用于将所述状态信息通过无线通信方式发送给所述拍摄设备，以使所述拍摄设备依据所述状态信息控制所述语音录制设备开启或关闭。

18. 根据权利要求 17 所述的语音录制设备，其特征在于，还包括：

AGC 限幅模块，用于对音量大于预设音量的语音数据进行 AGC 限幅处理；

所述语音录制设备的初始时钟与 GPS 时钟保持同步，所述拍摄设备的初始时钟与 GPS 时钟或所述拍摄设备所属基站的时钟保持同步。

19. 一种拍摄设备，其特征在于，包括：

发送模块，用于向语音录制设备发送时钟同步信息，所述时钟同步信息包括所述拍摄设备的起始拍摄时刻，以使所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻，并从所述起始拍摄时刻记录存储被拍摄者的语音数据，所述语音录制设备固定于所述被拍摄者；

接收模块，用于接收所述语音录制设备发送的所述语音数据；

合并模块，用于将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

20. 根据权利要求 19 所述的拍摄设备，其特征在于，所述发送模块还用于向语音录制设备发送初始时钟信息，所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳，以使所述语音录制设备依据所述时间戳开始计时。

21. 根据权利要求 20 所述的拍摄设备，其特征在于，所述语音数据和所述视频图像数据具有相同的时间轴；

所述合并模块具体用于依据所述相同的时间轴将所述语音数据和所述视频图像数据合并为音视频文件。

22. 根据权利要求 21 所述的拍摄设备，其特征在于，所述接收模块还用于接收所述语音录制设备发送的拍摄启动请求或拍摄结束请求；

所述拍摄设备还包括控制模块，用于依据所述拍摄启动请求启动拍摄功能或依据所述

拍摄结束请求关闭拍摄功能。

23. 根据权利要求 22 所述的拍摄设备, 其特征在于, 所述发送模块具体用于通过无线通信的方式向语音录制设备发送时钟同步信息;

所述接收模块具体用于通过无线通信的方式接收所述语音录制设备发送的所述语音数据。

24. 根据权利要求 23 所述的拍摄设备, 其特征在于, 所述接收模块还用于通过无线通信方式接收所述语音录制设备发送的状态信息;

所述控制模块还用于依据所述状态信息控制所述语音录制设备开启或关闭, 所述状态信息至少包括: 开关状态信息、电量状态信息和存储空间状态信息。

25. 一种音视频文件生成系统, 其特征在于, 包括如权利要求 13-18 任一项所述的语音录制设备以及如权利要求 19-24 任一项所述的拍摄设备。

音视频文件生成方法及装置

技术领域

[0001] 本发明实施例涉及通信技术领域，尤其涉及一种音视频文件生成方法及装置。

背景技术

[0002] 随着电子产品的发展，越来越多的电子产品具有拍摄功能，方便用户出行或随时随地拍摄。

[0003] 现有的具有拍摄功能的电子产品包括摄像机、平板电脑和手机等，在拍摄过程中，拍摄者需要手持电子产品即拍摄设备，通过拍摄设备的视频录制功能和语音录制功能同时记录下被拍摄者的画面和语音，即生成图文并茂的音视频文件。

[0004] 但是，当被拍摄者处于运动状态且远离拍摄者时，拍摄设备只能记录下被拍摄者的画面，却不能清晰记录被拍摄者的语音，导致录制的音视频文件中语音质量差。

发明内容

[0005] 本发明实施例提供一种音视频文件生成方法及装置，以提高音视频文件中语音质量。

[0006] 本发明实施例的一个方面是提供一种音视频文件生成方法，包括：

[0007] 语音录制设备接收拍摄设备发送的时钟同步信息，所述时钟同步信息包括所述拍摄设备的起始拍摄时刻；

[0008] 所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻，并从所述起始拍摄时刻记录存储被拍摄者的语音数据，所述语音录制设备固定于所述被拍摄者；

[0009] 所述语音录制设备将所述语音数据发送给所述拍摄设备，以使所述拍摄设备将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0010] 本发明实施例的另一个方面是提供一种音视频文件生成方法，包括：

[0011] 拍摄设备向语音录制设备发送时钟同步信息，所述时钟同步信息包括所述拍摄设备的起始拍摄时刻，以使所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻，并从所述起始拍摄时刻记录存储被拍摄者的语音数据，所述语音录制设备固定于所述被拍摄者；

[0012] 所述拍摄设备接收所述语音录制设备发送的所述语音数据，并将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0013] 本发明实施例的另一个方面是提供一种语音录制设备，包括：

[0014] 接收模块，用于接收拍摄设备发送的时钟同步信息，所述时钟同步信息包括所述拍摄设备的起始拍摄时刻；

[0015] 同步模块，用于设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻；

[0016] 记录存储模块,用于从所述起始拍摄时刻记录存储被拍摄者的语音数据,所述语音录制设备固定于所述被拍摄者;

[0017] 发送模块,用于将所述语音数据发送给所述拍摄设备,以使所述拍摄设备将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0018] 本发明实施例的另一个方面是提供一种拍摄设备,包括:

[0019] 发送模块,用于向语音录制设备发送时钟同步信息,所述时钟同步信息包括所述拍摄设备的起始拍摄时刻,以使所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻,并从所述起始拍摄时刻记录存储被拍摄者的语音数据,所述语音录制设备固定于所述被拍摄者;

[0020] 接收模块,用于接收所述语音录制设备发送的所述语音数据;

[0021] 合并模块,用于将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0022] 本发明实施例的另一个方面是提供一种音视频文件生成系统,包括所述的语音录制设备和所述的拍摄设备。

[0023] 本发明实施例提供的音视频文件生成方法及装置,通过语音录制设备接收拍摄设备的起始拍摄时刻,将其自身的时钟计时起点同步到起始拍摄时刻,并从起始拍摄时刻记录存储被拍摄者的语音数据,最终将该语音数据与拍摄设备从起始拍摄时刻拍摄的视频图像数据合并为音视频文件,由于语音数据和视频图像数据对应相同的起始拍摄时刻即基于相同的时间轴,即使被拍摄者处于运动状态且远离拍摄者时,固定于被拍摄者的语音录制设备能够记录存储清晰的语音数据,保证语音数据和视频图像数据具有很高的匹配度,并保证合并后的音视频文件具有很高的语音质量。

附图说明

[0024] 图 1 为本发明实施例提供的音视频文件生成方法流程图;

[0025] 图 2 为本发明实施例提供的音视频文件生成方法流程图;

[0026] 图 3 为本发明实施例提供的语音录制设备的结构图;

[0027] 图 4 为本发明另一实施例提供的语音录制设备的结构图;

[0028] 图 5 为本发明实施例提供的拍摄设备的结构图;

[0029] 图 6 为本发明另一实施例提供的拍摄设备的结构图;

[0030] 图 7 为本发明实施例提供的音视频文件生成系统的结构图。

具体实施方式

[0031] 图 1 为本发明实施例提供的音视频文件生成方法流程图。本发明实施例中的语音录制设备具体为一个蓝夹子结构,该蓝夹子结构包括夹子和话筒,夹子用于固定在被拍摄者的衣服上便于录制被拍摄者的语音数据,本发明实施例优选夹持性很强的夹子固定在被拍摄者的衣领上。话筒与夹子固定连接,话筒内部包括存储模块、录音模块、AGC 限幅模块、时钟模块(晶振)、全球定位系统(全球定位系统,简称 GPS)模块以及无线通信模块(wifi、蓝牙)等,话筒通过无线通信模块与拍摄设备进行无线连接,该拍摄设备具体为智能手机、

相机等。

[0032] 本发明实施例针对被拍摄者处于运动状态且远离拍摄者时，拍摄设备只能记录下被拍摄者的画面，却不能清晰记录被拍摄者的语音，提供了音视频文件生成方法，该方法具体步骤如下：

[0033] 步骤 S101、语音录制设备接收拍摄设备发送的时钟同步信息，所述时钟同步信息包括所述拍摄设备的起始拍摄时刻；

[0034] 拍摄设备例如智能手机在开始拍摄视频时刻，向语音录制设备即话筒发送时钟同步信息，该时钟同步信息中包含有智能手机的起始拍摄时刻。

[0035] 步骤 S102、所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻，并从所述起始拍摄时刻记录存储被拍摄者的语音数据，所述语音录制设备固定于所述被拍摄者；

[0036] 话筒在接收拍摄设备发送的时钟同步信息之前，其时钟计时起点与 GPS 时钟保持同步，并具有很高的走时精度，当话筒接收到拍摄设备发送的时钟同步信息时，话筒依据所述时钟同步信息将其自身的时钟计时起点同步到所述起始拍摄时刻，并从所述起始拍摄时刻起记录被拍摄者的语音数据，记录的语音数据存储在话筒的存储模块中。

[0037] 步骤 S103、所述语音录制设备将所述语音数据发送给所述拍摄设备，以使所述拍摄设备将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0038] 话筒可以将其存储的语音数据实时发送给拍摄设备，也可以按固定时间间隔发送给拍摄设备，还可以在拍摄设备拍摄视频结束后发送给拍摄设备，本发明实施例不做限制，另外，话筒向拍摄设备发送语音数据的方式优选为无线通信方式，具体为无线高保真(Wireless Fidelity，简称 WiFi)、蓝牙等。拍摄设备接收到话筒发送的语音数据后，将语音数据和拍摄设备从起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件，由于语音数据和视频图像数据对应相同的起始拍摄时刻，所以语音数据和视频图像数据具有很高的匹配度，即合并后的音视频文件同时具有高质量的视觉与听觉效果。

[0039] 另外，将语音数据和视频图像数据合并为音视频文件的过程还可以在计算机上完成，具体为语音数据和视频图像数据均录制结束后，用户将话筒中的语音数据和拍摄设备中的视频图像数据分别拷贝到计算机上，由计算机将语音数据和视频图像数据合并为音视频文件。

[0040] 本发明实施例通过语音录制设备接收拍摄设备的起始拍摄时刻，将其自身的时钟计时起点同步到起始拍摄时刻，并从起始拍摄时刻记录存储被拍摄者的语音数据，最终将该语音数据与拍摄设备从起始拍摄时刻拍摄的视频图像数据合并为音视频文件，由于语音数据和视频图像数据对应相同的起始拍摄时刻即基于相同的时间轴，即使被拍摄者处于运动状态且远离拍摄者时，固定于被拍摄者的语音录制设备能够记录存储清晰的语音数据，保证语音数据和视频图像数据具有很高的匹配度，并保证合并后的音视频文件具有很高的语音质量。

[0041] 在上述实施例的基础上，所述语音录制设备接收拍摄设备发送的时钟同步信息之前，还包括：所述语音录制设备接收所述拍摄设备发送的初始时钟信息，所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳；所述语音录制设备依据所述

时间戳开始计时。

[0042] 在本发明实施例中，所述语音录制设备与所述拍摄设备配对出现，即所述语音录制设备接收拍摄设备发送的时钟同步信息之前，需要对所述语音录制设备与所述拍摄设备进行配对处理，具体的配对处理过程为：拍摄设备向语音录制设备发送初始时钟信息，初始时钟信息包括的一个时间戳，该时间戳是拍摄设备发送所述初始时钟信息时刻对应的时间戳。语音录制设备依据所述时间戳开始计时，即语音录制设备的计时开始时刻以拍摄设备发送初始时钟信息的时刻为准。

[0043] 所述语音录制设备依据所述时间戳开始计时之后，还包括：所述语音录制设备向所述拍摄设备发送拍摄启动请求，以使所述拍摄设备启动拍摄功能；所述语音录制设备将所述语音数据发送给所述拍摄设备之后，还包括：所述语音录制设备向所述拍摄设备发送拍摄结束请求，以使所述拍摄设备关闭拍摄功能。

[0044] 本发明实施例中的语音录制设备和拍摄设备还可以应用在水下拍摄场景中。语音录制设备可以作为主控设备控制拍摄设备对应拍摄功能的开启或关闭，具体为语音录制设备先向拍摄设备发送拍摄启动请求，以使所述拍摄设备启动拍摄功能，拍摄设备将其起始拍摄时刻发送给语音录制设备；当语音录制设备记录语音数据结束后，主动向拍摄设备发送拍摄结束请求，拍摄设备依据该拍摄结束请求关闭拍摄功能。

[0045] 所述语音录制设备接收拍摄设备发送的时钟同步信息，包括：所述语音录制设备通过无线通信的方式接收拍摄设备发送的时钟同步信息；所述语音录制设备将所述语音数据发送给所述拍摄设备，包括：所述语音录制设备将所述语音数据通过无线通信的方式发送给所述拍摄设备。

[0046] 在本发明实施例中，语音录制设备与拍摄设备之间的所有交互均采用无线通信的方式进行数据传输。

[0047] 本发明实施例通过语音录制设备接收拍摄设备发送的时钟同步信息之前接收拍摄设备发送的初始时钟信息，并依据初始时钟信息中的时间戳开始计时，防止语音录制设备接收不到时钟同步信息时，无法与拍摄设备保持时钟同步，进一步提高了时钟同步精度；另外通过语音录制设备作为主控设备控制拍摄设备对应拍摄功能的开启或关闭，增加了语音录制设备的功能。

[0048] 在上述实施例的基础上，所述语音录制设备监测其自身的开关状态、电量状态和存储空间状态，并生成状态信息，所述状态信息至少包括：开关状态信息、电量状态信息和存储空间状态信息；所述语音录制设备将所述状态信息通过无线通信方式发送给所述拍摄设备，以使所述拍摄设备依据所述状态信息控制所述语音录制设备开启或关闭。

[0049] 本发明实施例中的语音录制设备还设置有物理的开关按钮、开关指示灯、电量指示灯、容量指示灯，开关指示灯用于指示语音录制设备的开关状态，电量指示灯用于指示语音录制设备的电量状态，容量指示灯用于指示语音录制设备的存储空间状态。语音录制设备监测其自身的开关状态、电量状态和存储空间状态，并生成相应状态信息，同时通过开关指示灯指示开关状态，若电量低于电量阈值时，通过电量指示灯进行告警显示，当存储空间小于预设阈值时，通过开关指示灯进行告警显示。

[0050] 语音录制设备将监测到的状态信息通过无线通信方式发送给拍摄设备，拍摄设备作为主控设备依据语音录制设备的状态信息控制语音录制设备开启或关闭，例如语音录制

设备的电量低于电量阈值或存储空间小于预设阈值时,拍摄设备通过无线通信方式向语音录制设备发送关闭指令,以使语音录制设备依据该关闭指令执行关闭操作。

[0051] 本发明实施例通过语音录制设备将其状态信息发送给拍摄设备,由拍摄设备依据该状态信息控制语音录制设备开启或关闭,提高了拍摄设备对语音录制设备的控制功能。

[0052] 在上述实施例的基础上,所述从所述起始拍摄时刻记录存储被拍摄者的语音数据之后,还包括:所述语音录制设备对音量大于预设音量的语音数据进行 AGC 限幅处理;所述语音录制设备的初始时钟与 GPS 时钟保持同步,所述拍摄设备的初始时钟与 GPS 时钟或所述拍摄设备所属基站的时钟保持同步。

[0053] 语音录制设备还包括 AGC 限幅器,当语音录制设备录制的语音数据的音量大于预设音量时,AGC 限幅器对该语音数据的音量进行 AGC 限幅处理。另外,所述语音录制设备的初始时钟选取 GPS 时钟,所述拍摄设备的初始时钟选取 GPS 时钟或所述拍摄设备所属基站的时钟。

[0054] 另外,本发明实施例中的语音录制设备上还可设置一个比如 3.5mm 接口,可以适配所有市面上的有线耳机和话筒,从而无需单独携带一个额外的话筒,将其与现有耳机、话筒集成设置在一起即可。

[0055] 此外,语音录制设备还可以作为独立的录音设备使用即不需与拍摄设备配合使用。

[0056] 本发明实施例通过语音录制设备对音量大于预设音量的语音数据进行 AGC 限幅处理,防止语音录制设备的录音模块被较大的语音数据震坏,进一步保证了合并后的音视频文件具有很高的语音质量。

[0057] 图 2 为本发明实施例提供的音视频文件生成方法流程图,本发明实施例针对被拍摄者处于运动状态且远离拍摄者时,拍摄设备只能记录下被拍摄者的画面,却不能清晰记录被拍摄者的语音,提供了音视频文件生成方法,该方法具体步骤如下:

[0058] 步骤 S201、拍摄设备向语音录制设备发送时钟同步信息,所述时钟同步信息包括所述拍摄设备的起始拍摄时刻,以使所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻,并从所述起始拍摄时刻记录存储被拍摄者的语音数据,所述语音录制设备固定于所述被拍摄者;

[0059] 拍摄设备例如智能手机在开始拍摄视频时刻,向语音录制设备即话筒发送时钟同步信息,该时钟同步信息中含有智能手机的起始拍摄时刻。话筒在接收拍摄设备发送的时钟同步信息之前,其时钟计时起点与 GPS 时钟保持同步,并具有很高的走时精度,当话筒接收到拍摄设备发送的时钟同步信息时,话筒依据所述时钟同步信息将其自身的时钟计时起点同步到所述起始拍摄时刻,并从所述起始拍摄时刻起记录被拍摄者的语音数据,记录的语音数据存储在话筒的存储模块中。

[0060] 步骤 S202、所述拍摄设备接收所述语音录制设备发送的所述语音数据,并将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0061] 话筒可以将其存储的语音数据实时发送给拍摄设备,也可以按固定时间间隔发送给拍摄设备,还可以在拍摄设备拍摄视频结束后发送给拍摄设备,本发明实施例不做限制,另外,话筒向拍摄设备发送语音数据的方式优选为无线通信方式,具体为无线高保真

(Wireless Fidelity, 简称 WiFi)、蓝牙等。拍摄设备接收到话筒发送的语音数据后, 将语音数据和拍摄设备从起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件, 由于语音数据和视频图像数据对应相同的起始拍摄时刻, 所以语音数据和视频图像数据具有很高的匹配度, 即合并后的音视频文件同时具有高质量的视觉与听觉效果。

[0062] 本发明实施例通过语音录制设备接收拍摄设备的起始拍摄时刻, 将其自身的时钟计时起点同步到起始拍摄时刻, 并从起始拍摄时刻记录存储被拍摄者的语音数据, 最终将该语音数据与拍摄设备从起始拍摄时刻拍摄的视频图像数据合并为音视频文件, 由于语音数据和视频图像数据对应相同的起始拍摄时刻即基于相同的时间轴, 即使被拍摄者处于运动状态且远离拍摄者时, 固定于被拍摄者的语音录制设备能够记录存储清晰的语音数据, 保证语音数据和视频图像数据具有很高的匹配度, 并保证合并后的音视频文件具有很高的语音质量。

[0063] 在上述实施例的基础上, 所述拍摄设备向语音录制设备发送时钟同步信息之前, 还包括: 所述拍摄设备向语音录制设备发送初始时钟信息, 所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳, 以使所述语音录制设备依据所述时间戳开始计时。

[0064] 在本发明实施例中, 所述语音录制设备与所述拍摄设备配对出现, 即所述语音录制设备接收拍摄设备发送的时钟同步信息之前, 需要对所述语音录制设备与所述拍摄设备进行配对处理, 具体的配对处理过程为: 拍摄设备向语音录制设备发送初始时钟信息, 初始时钟信息包括的一个时间戳, 该时间戳是拍摄设备发送所述初始时钟信息时刻对应的时间戳。语音录制设备依据所述时间戳开始计时, 即语音录制设备的计时开始时刻以拍摄设备发送初始时钟信息的时刻为准。

[0065] 所述语音数据和所述视频图像数据具有相同的时间轴; 所述将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件, 包括: 依据所述相同的时间轴将所述语音数据和所述视频图像数据合并为音视频文件。

[0066] 所述拍摄设备向语音录制设备发送时钟同步信息之前, 还包括: 所述拍摄设备接收所述语音录制设备发送的拍摄启动请求, 并依据所述拍摄启动请求启动拍摄功能; 所述拍摄设备接收所述语音录制设备发送的所述语音数据之后, 还包括: 所述拍摄设备接收所述语音录制设备发送的拍摄结束请求, 并依据所述拍摄结束请求关闭拍摄功能。

[0067] 语音录制设备可以作为主控设备控制拍摄设备对应拍摄功能的开启或关闭, 具体为语音录制设备先向拍摄设备发送拍摄启动请求, 以使所述拍摄设备启动拍摄功能, 拍摄设备将其起始拍摄时刻发送给语音录制设备; 当语音录制设备记录语音数据结束后, 主动向拍摄设备发送拍摄结束请求, 拍摄设备依据该拍摄结束请求关闭拍摄功能。

[0068] 所述拍摄设备向语音录制设备发送时钟同步信息, 包括: 所述拍摄设备通过无线通信的方式向语音录制设备发送时钟同步信息; 所述拍摄设备接收所述语音录制设备发送的所述语音数据, 包括: 所述拍摄设备通过无线通信的方式接收所述语音录制设备发送的所述语音数据。

[0069] 在本发明实施例中, 语音录制设备与拍摄设备之间的所有交互均采用无线通信的方式进行数据传输。

[0070] 所述拍摄设备通过无线通信方式接收所述语音录制设备发送的状态信息, 并依据

所述状态信息控制所述语音录制设备开启或关闭,所述状态信息至少包括:开关状态信息、电量状态信息和存储空间状态信息。

[0071] 本发明实施例中的语音录制设备还设置有物理的开关按钮、开关指示灯、电量指示灯、容量指示灯,开关指示灯用于指示语音录制设备的开关状态,电量指示灯用于指示语音录制设备的电量状态,容量指示灯用于指示语音录制设备的存储空间状态。语音录制设备监测其自身的开关状态、电量状态和存储空间状态,并生成相应状态信息,同时通过开关指示灯指示开关状态,若电量低于电量阈值时,通过电量指示灯进行告警显示,当存储空间小于预设阈值时,通过开关指示灯进行告警显示。

[0072] 语音录制设备将监测到的状态信息通过无线通信方式发送给拍摄设备,拍摄设备作为主控设备依据语音录制设备的状态信息控制语音录制设备开启或关闭,例如语音录制设备的电量低于电量阈值或存储空间小于预设阈值时,拍摄设备通过无线通信方式向语音录制设备发送关闭指令,以使语音录制设备依据该关闭指令执行关闭操作。

[0073] 本发明实施例通过语音录制设备接收拍摄设备发送的时钟同步信息之前接收拍摄设备发送的初始时钟信息,并依据初始时钟信息中的时间戳开始计时,防止语音录制设备接收不到时钟同步信息时,无法与拍摄设备保持时钟同步,进一步提高了时钟同步精度;通过语音录制设备作为主控设备控制拍摄设备对应拍摄功能的开启或关闭,增加了语音录制设备的功能;通过语音录制设备将其状态信息发送给拍摄设备,由拍摄设备依据该状态信息控制语音录制设备开启或关闭,提高了拍摄设备对语音录制设备的控制功能。

[0074] 图3为本发明实施例提供的语音录制设备的结构图。本发明实施例提供的语音录制设备可以执行音视频文件生成方法实施例提供的处理流程,如图3所示,语音录制设备30包括接收模块31、同步模块32、记录存储模块33和发送模块34,其中,接收模块31用于接收拍摄设备发送的时钟同步信息,所述时钟同步信息包括所述拍摄设备的起始拍摄时刻;同步模块32用于设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻;记录存储模块33用于从所述起始拍摄时刻记录存储被拍摄者的语音数据,所述语音录制设备固定于所述被拍摄者;发送模块34用于将所述语音数据发送给所述拍摄设备,以使所述拍摄设备将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0075] 本发明实施例通过语音录制设备接收拍摄设备的起始拍摄时刻,将其自身的时钟计时起点同步到起始拍摄时刻,并从起始拍摄时刻记录存储被拍摄者的语音数据,最终将该语音数据与拍摄设备从起始拍摄时刻拍摄的视频图像数据合并为音视频文件,由于语音数据和视频图像数据对应相同的起始拍摄时刻即基于相同的时间轴,即使被拍摄者处于运动状态且远离拍摄者时,固定于被拍摄者的语音录制设备能够记录存储清晰的语音数据,保证语音数据和视频图像数据具有很高的匹配度,并保证合并后的音视频文件具有很高的语音质量。

[0076] 图4为本发明另一实施例提供的语音录制设备的结构图。在上述实施例的基础上,接收模块31还用于接收所述拍摄设备发送的初始时钟信息,所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳;

[0077] 语音录制设备30还包括计时模块37,时模块37用于依据所述时间戳开始计时。

[0078] 发送模块34还用于所述语音录制设备依据所述时间戳开始计时之后,向所述拍

摄设备发送拍摄启动请求,以使所述拍摄设备启动拍摄功能;所述语音录制设备将所述语音数据发送给所述拍摄设备之后,向所述拍摄设备发送拍摄结束请求,以使所述拍摄设备关闭拍摄功能。

[0079] 接收模块 31 具体用于通过无线通信的方式接收拍摄设备发送的时钟同步信息;发送模块 34 具体用于将所述语音数据通过无线通信的方式发送给所述拍摄设备。

[0080] 语音录制设备 30 还包括监测模块 35,监测模块 35 用于监测其自身的开关状态、电量状态和存储空间状态,并生成状态信息,所述状态信息至少包括:开关状态信息、电量状态信息和存储空间状态信息;发送模块 34 还用于将所述状态信息通过无线通信方式发送给所述拍摄设备,以使所述拍摄设备依据所述状态信息控制所述语音录制设备开启或关闭。

[0081] 语音录制设备 30 还包括 AGC 限幅模块 36,AGC 限幅模块 36 用于对音量大于预设音量的语音数据进行 AGC 限幅处理;所述语音录制设备的初始时钟与 GPS 时钟保持同步,所述拍摄设备的初始时钟与 GPS 时钟或所述拍摄设备所属基站的时钟保持同步。

[0082] 本发明实施例提供的语音录制设备可以具体用于执行上述图 1 所提供的方法实施例,具体功能此处不再赘述。

[0083] 本发明实施例通过语音录制设备接收拍摄设备发送的时钟同步信息之前接收拍摄设备发送的初始时钟信息,并依据初始时钟信息中的时间戳开始计时,防止语音录制设备接收不到时钟同步信息时,无法与拍摄设备保持时钟同步,进一步提高了时钟同步精度;通过语音录制设备作为主控设备控制拍摄设备对应拍摄功能的开启或关闭,增加了语音录制设备的功能;通过语音录制设备将其状态信息发送给拍摄设备,由拍摄设备依据该状态信息控制语音录制设备开启或关闭,提高了拍摄设备对语音录制设备的控制功能;通过语音录制设备对音量大于预设音量的语音数据进行 AGC 限幅处理,防止语音录制设备的录音模块被较大的语音数据震坏,进一步保证了合并后的音视频文件具有很高的语音质量。

[0084] 图 5 为本发明实施例提供的拍摄设备的结构图。本发明实施例提供的拍摄设备可以执行音视频文件生成方法实施例提供的处理流程,如图 5 所示,拍摄设备 50 包括发送模块 51、接收模块 52 和合并模块 53,其中,发送模块 51 用于向语音录制设备发送时钟同步信息,所述时钟同步信息包括所述拍摄设备的起始拍摄时刻,以使所述语音录制设备依据所述时钟同步信息将所述语音录制设备的时钟计时起点同步到所述起始拍摄时刻,并从所述起始拍摄时刻记录存储被拍摄者的语音数据,所述语音录制设备固定于所述被拍摄者;接收模块 52 用于接收所述语音录制设备发送的所述语音数据;合并模块 53 用于将所述语音数据和所述拍摄设备从所述起始拍摄时刻拍摄的被拍摄者的视频图像数据合并为音视频文件。

[0085] 本发明实施例通过语音录制设备接收拍摄设备的起始拍摄时刻,将其自身的时钟计时起点同步到起始拍摄时刻,并从起始拍摄时刻记录存储被拍摄者的语音数据,最终将该语音数据与拍摄设备从起始拍摄时刻拍摄的视频图像数据合并为音视频文件,由于语音数据和视频图像数据对应相同的起始拍摄时刻即基于相同的时间轴,即使被拍摄者处于运动状态且远离拍摄者时,固定于被拍摄者的语音录制设备能够记录存储清晰的语音数据,保证语音数据和视频图像数据具有很高的匹配度,并保证合并后的音视频文件具有很高的语音质量。

[0086] 图 6 为本发明另一实施例提供的拍摄设备的结构图。在上述实施例的基础上，发送模块 51 还用于向语音录制设备发送初始时钟信息，所述初始时钟信息包括所述拍摄设备发送所述初始时钟信息时刻对应的时间戳，以使所述语音录制设备依据所述时间戳开始计时。

[0087] 所述语音数据和所述视频图像数据具有相同的时间轴；合并模块 53 具体用于依据所述相同的时间轴将所述语音数据和所述视频图像数据合并为音视频文件。

[0088] 接收模块 52 还用于接收所述语音录制设备发送的拍摄启动请求或拍摄结束请求；拍摄设备 50 还包括控制模块 54，控制模块 54 用于依据所述拍摄启动请求启动拍摄功能或依据所述拍摄结束请求关闭拍摄功能。

[0089] 发送模块 51 具体用于通过无线通信的方式向语音录制设备发送时钟同步信息；接收模块 52 具体用于通过无线通信的方式接收所述语音录制设备发送的所述语音数据。

[0090] 接收模块 52 还用于通过无线通信方式接收所述语音录制设备发送的状态信息；控制模块 54 还用于依据所述状态信息控制所述语音录制设备开启或关闭，所述状态信息至少包括：开关状态信息、电量状态信息和存储空间状态信息。

[0091] 本发明实施例提供的拍摄设备可以具体用于执行上述图 2 所提供的方法实施例，具体功能此处不再赘述。

[0092] 本发明实施例通过语音录制设备接收拍摄设备发送的时钟同步信息之前接收拍摄设备发送的初始时钟信息，并依据初始时钟信息中的时间戳开始计时，防止语音录制设备接收不到时钟同步信息时，无法与拍摄设备保持时钟同步，进一步提高了时钟同步精度；通过语音录制设备作为主控设备控制拍摄设备对应拍摄功能的开启或关闭，增加了语音录制设备的功能；通过语音录制设备将其状态信息发送给拍摄设备，由拍摄设备依据该状态信息控制语音录制设备开启或关闭，提高了拍摄设备对语音录制设备的控制功能。

[0093] 图 7 为本发明实施例提供的音视频文件生成系统的结构图。本发明实施例提供的音视频文件生成系统可以执行音视频文件生成方法实施例提供的处理流程，如图 7 所示，音视频文件生成系统 70 包括上述实施例中的语音录制设备 30 和上述实施例中的拍摄设备 50。

[0094] 本发明实施例提供的音视频文件生成系统可以执行音视频文件生成方法实施例提供的处理流程。

[0095] 综上所述，本发明实施例通过语音录制设备接收拍摄设备的起始拍摄时刻，将其自身的时钟计时起点同步到起始拍摄时刻，并从起始拍摄时刻记录存储被拍摄者的语音数据，最终将该语音数据与拍摄设备从起始拍摄时刻拍摄的视频图像数据合并为音视频文件，由于语音数据和视频图像数据对应相同的起始拍摄时刻即基于相同的时间轴，即使被拍摄者处于运动状态且远离拍摄者时，固定于被拍摄者的语音录制设备能够记录存储清晰的语音数据，保证语音数据和视频图像数据具有很高的匹配度，并保证合并后的音视频文件具有很高的语音质量；通过语音录制设备接收拍摄设备发送的时钟同步信息之前接收拍摄设备发送的初始时钟信息，并依据初始时钟信息中的时间戳开始计时，防止语音录制设备接收不到时钟同步信息时，无法与拍摄设备保持时钟同步，进一步提高了时钟同步精度；通过语音录制设备作为主控设备控制拍摄设备对应拍摄功能的开启或关闭，增加了语音录制设备的功能；通过语音录制设备将其状态信息发送给拍摄设备，由拍摄设备依据该状态

信息控制语音录制设备开启或关闭，提高了拍摄设备对语音录制设备的控制功能；通过语音录制设备对音量大于预设音量的语音数据进行 AGC 限幅处理，防止语音录制设备的录音模块被较大的语音数据震坏，进一步保证了合并后的音视频文件具有很高的语音质量。

[0096] 在本发明所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，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。

[0097] 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例方案的目的。

[0098] 另外，在本发明各个实施例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，也可以采用硬件加软件功能单元的形式实现。

[0099] 上述以软件功能单元的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能单元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例所述方法的部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory，ROM）、随机存取存储器（Random Access Memory，RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

[0100] 本领域技术人员可以清楚地了解到，为描述的方便和简洁，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。上述描述的装置的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

[0101] 最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

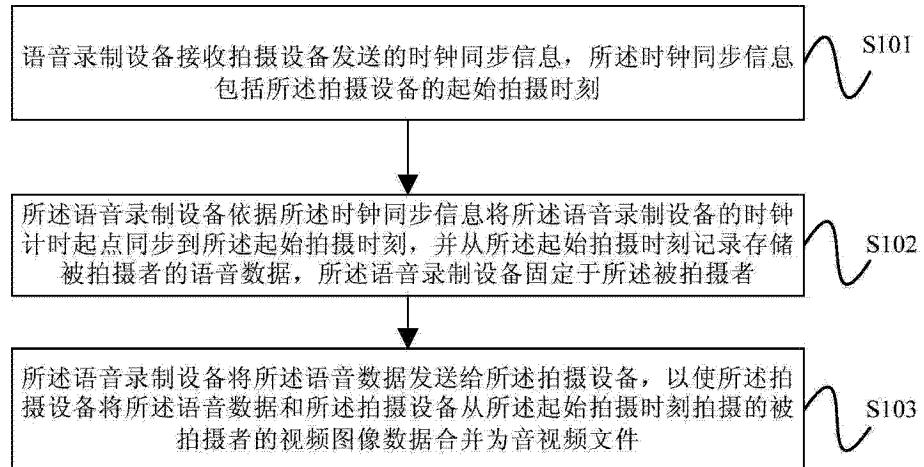


图 1

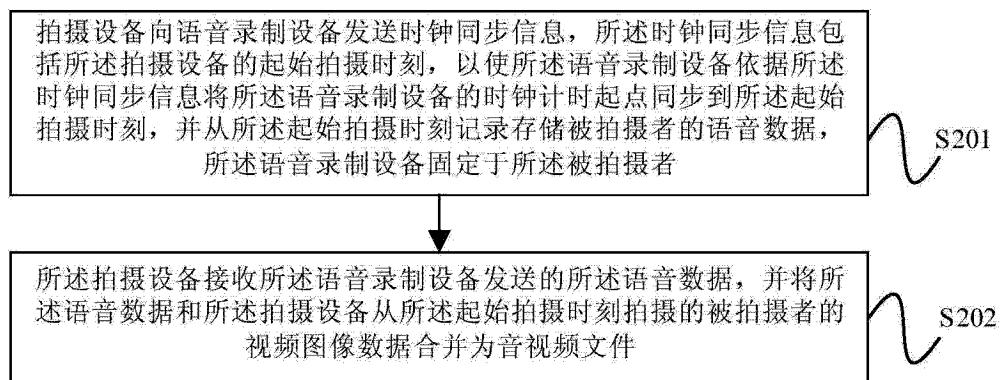


图 2

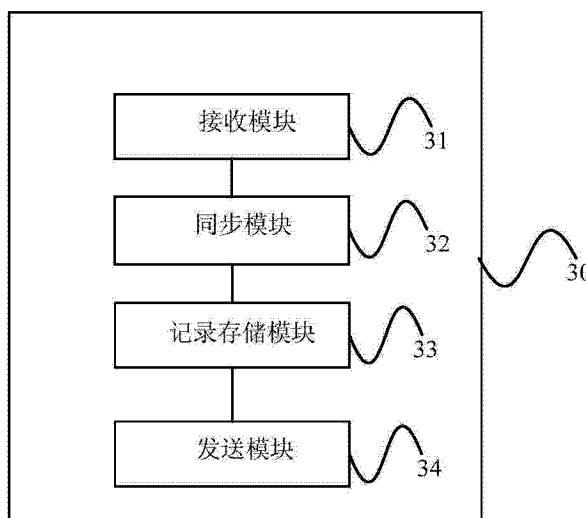


图 3

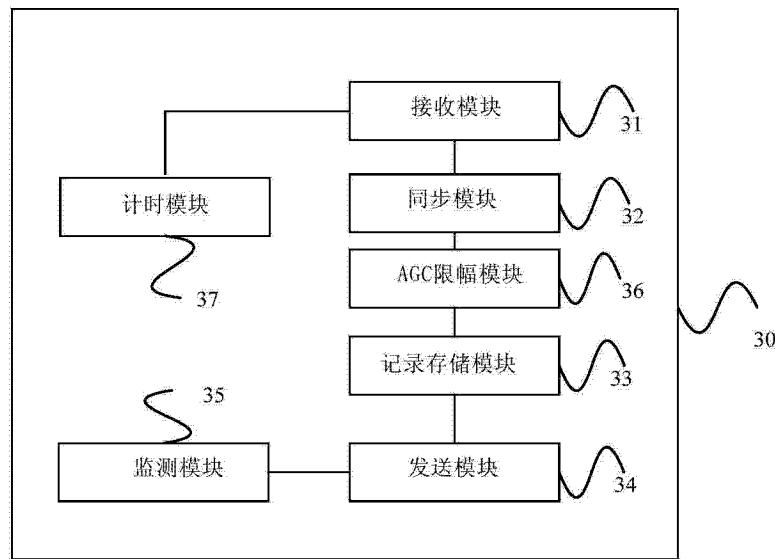


图 4

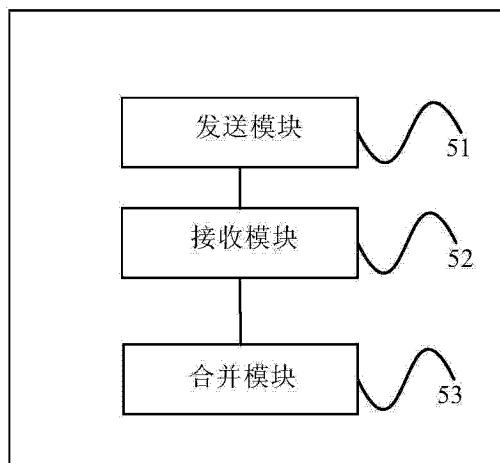


图 5

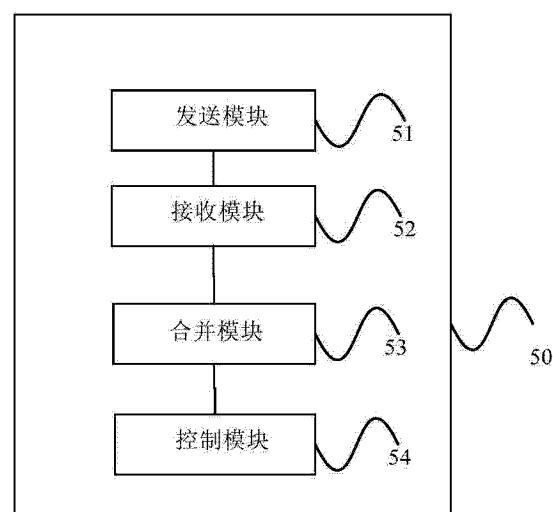


图 6

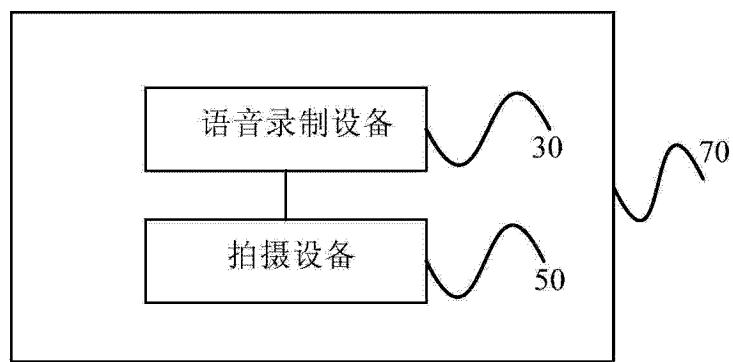


图 7