



(12) 发明专利申请

(10) 申请公布号 CN 111918077 A

(43) 申请公布日 2020.11.10

(21) 申请号 202010718614.9

H04N 21/4402 (2011.01)

(22) 申请日 2020.07.23

(71) 申请人 安徽声讯信息技术有限公司

地址 230000 安徽省合肥市合肥高新技术
开发区习友路与习南莲路交口中国声
谷

(72) 发明人 虞焰兴

(74) 专利代理机构 上海精晟知识产权代理有限
公司 31253

代理人 李佼佼

(51) Int. Cl.

H04N 21/2187 (2011.01)

H04N 21/2368 (2011.01)

H04N 21/242 (2011.01)

H04N 21/431 (2011.01)

权利要求书1页 说明书3页 附图1页

(54) 发明名称

视频直播图像声音采集和电脑桌面采集同
步合成方法

(57) 摘要

本发明公开了一种视频直播图像声音采集
和电脑桌面采集同步合成方法,涉及流媒体合成
技术领域。本发明包括如下步骤:获取在桌面预
选的至少一个显示区域的音频/位置参数;抓取
当前整个桌面视频信息进行绑定为第一数据包;
摄像头采集教师上课视频进行绑定为第二数据
包;将第一数据包和第二数据包进行解析合并;
合并后的视频发送给服务器端进行区域音频/位
置参数集的音频数据和视频图像进行匹配;匹配
成功后的视频数据在指定区域位置进行覆盖。本
发明方便学生远程观看授课的课件和老师的同
步讲解,同时多个视频框随意切换,提高学生远
程上课的学习效果。



1. 一种视频直播图像声音采集和电脑桌面采集同步合成方法,其特征在于,包括如下步骤:

步骤S1:获取在桌面预选的至少一个显示区域的音频/位置参数;

步骤S2:抓取当前整个桌面视频信息进行绑定为第一数据包;

步骤S3:摄像头采集教师上课视频进行绑定为第二数据包;

步骤S4:将第一数据包和第二数据包进行解析合并;

步骤S5:合并后的视频发送给服务器端进行区域音频/位置参数集的音频数据和视频图像进行匹配;

步骤S6:匹配成功后的视频数据在指定区域位置进行覆盖。

2. 根据权利要求1所述的一种视频直播图像声音采集和电脑桌面采集同步合成方法,其特征在于,所述步骤S1中,桌面视频和摄像头采集的上课视频可以在显示区域内进行切换。

3. 根据权利要求1所述的一种视频直播图像声音采集和电脑桌面采集同步合成方法,其特征在于,所述步骤S2中,客户端截取整个桌面的视频图像,并进行编码压缩获取视频编码数据包;所述视频图像包括至少一个显示区域;所述视频编码数据包与至少一个所述位置参数绑定为对应的第二网络数据包,并发送给所述目标服务端。

4. 根据权利要求1所述的一种视频直播图像声音采集和电脑桌面采集同步合成方法,其特征在于,所述步骤S4中,对桌面视频进行音频标志的标识处理,并形成区域音频/位置参数集,并将提取并编码压缩的音频数据与区域音频/位置参数集进行绑定为对应的第一网络数据包,并发送给所述目标服务端。

5. 根据权利要求1或4所述的一种视频直播图像声音采集和电脑桌面采集同步合成方法,其特征在于,所述目标服务端对或去的第一网络数据包或第二网络数据包进行类型解析,若为第一网络数据包,则对编码压缩后的音频数据进行解码获取至少一组音频数据,若为所述第二网络数据包,则对编码压缩后的音频数据进行解码获取到视频图像。

6. 根据权利要求4所述的一种视频直播图像声音采集和电脑桌面采集同步合成方法,其特征在于,所述目标服务端对位置参数对应的显示区域进行是否存在有音频标志的检测,若有,则根据音频标志对应的区域音频/位置参数集进行音频数据和位置参数对应的视频图像的匹配,并进行对应的客户端桌面区域共享,若否,则位置参数对应的视频图像进行对应的客户端桌面区域共享。

7. 根据权利要求1所述的一种视频直播图像声音采集和电脑桌面采集同步合成方法,其特征在于,所述S5中,若第一数据包和第二数据包进行解析合并使均包含音频标志,则需要对合并后的音频数据进行混音处理。

视频直播图像声音采集和电脑桌面采集同步合成方法

技术领域

[0001] 本发明属于流媒体合成技术领域,特别是涉及一种视频直播图像声音采集和电脑桌面采集同步合成方法。

背景技术

[0002] 多媒体(Multimedia)是多种媒体的综合,一般包括文本、声音和图像等多种媒体形式。多媒体是信息化的体现,也是社会发展的趋势,多媒体教育是信息化的一个部分,大力促进多媒体教育已成为教育发展的潮流。

[0003] 由于多媒体固有的融合文本、声音和图像等多种媒体形式的特点,教学中往往使用多媒体技术,有利于提高教师的专业水平,有利于教师整合教学资源。多媒体教学技术能弥补传统教学中的不足,是新型的科学教学技术。

[0004] 目前,在采用多媒体技术的远程培训中,讲师需要将课件(PPT/或其他文档)、电脑桌面向学生共享,但上课学生看电脑的时候看不到教师的表情、动作、肢体语言,因此会遗漏老师的肢体表达信息,除此之外,其互动性及趣味性也较差;而且在直播的过程中,由于流媒体的数据帧非常容易丢帧,导致视频卡顿或者声音、画面不同步的问题,严重影响教师的教学质量。

发明内容

[0005] 本发明的目的在于提供一种视频直播图像声音采集和电脑桌面采集同步合成方法,通过将桌面视频信息和教师上课视频信息进行采集解析合并,并将合并后的视频再服务器进行区域匹配,匹配成功后在指定区域进行覆盖,解决了现有的远程培训课堂老师和电脑桌面不能同步、授课效率低、直播视频容易卡顿的问题。

[0006] 为解决上述技术问题,本发明是通过以下技术方案实现的:

[0007] 本发明为一种视频直播图像声音采集和电脑桌面采集同步合成方法,包括如下步骤:

[0008] 步骤S1:获取在桌面预选的至少一个显示区域的音频/位置参数;

[0009] 步骤S2:抓取当前整个桌面视频信息进行绑定为第一数据包;

[0010] 步骤S3:摄像头采集教师上课视频进行绑定为第二数据包;

[0011] 步骤S4:将第一数据包和第二数据包进行解析合并;

[0012] 步骤S5:合并后的视频发送给服务器端进行区域音频/位置参数集的音频数据和视频图像进行匹配;

[0013] 步骤S6:匹配成功后的视频数据在指定区域位置进行覆盖。

[0014] 优选地,所述步骤S1中,桌面视频和摄像头采集的上课视频可以在显示区域内进行切换。

[0015] 优选地,所述步骤S2中,客户端截取整个桌面的视频图像,并进行编码压缩获取视频编码数据包;所述视频图像包括至少一个显示区域;所述视频编码数据包与至少一个所

述位置参数绑定为对应的第二网络数据包,并发送给所述目标服务端。

[0016] 优选地,所述步骤S4中,对桌面视频进行音频标志的标识处理,并形成区域音频/位置参数集,并将提取并编码压缩的音频数据与区域音频/位置参数集进行绑定为对应的第一网络数据包,并发送给所述目标服务端。

[0017] 优选地,所述目标服务端对或去的第一网络数据包或第二网络数据包进行类型解析,若为第一网络数据包,则对编码压缩后的音频数据进行解码获取至少一组音频数据,若为所述第二网络数据包,则对编码压缩后的音频数据进行解码获取到视频图像。

[0018] 优选地,所述目标服务端对位置参数对应的显示区域进行是否存在有音频标志的检测,若有,则根据音频标志对应的区域音频/位置参数集进行音频数据和位置参数对应的视频图像的匹配,并进行对应的客户端桌面区域共享,若无,则位置参数对应的视频图像进行对应的客户端桌面区域共享。

[0019] 优选地,所述S5中,若第一数据包和第二数据包进行解析合并使均包含音频标志,则需要对合并后的音频数据进行混音处理。

[0020] 本发明具有以下有益效果:

[0021] 本发明通过将桌面视频信息和教师上课视频信息进行采集解析合并,并将合并后的视频再服务器进行区域匹配,匹配成功后在指定区域进行覆盖,方便学生远程观看授课的课件和老师的同步讲解,同时多个视频框随意切换,方便学生进行观摩学习,提高学生远程上课的学习效果。

[0022] 当然,实施本发明的任一产品并不一定需要同时达到以上所述的所有优点。

附图说明

[0023] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例描述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发明的一些实施例,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。

[0024] 图1为本发明的一种视频直播图像声音采集和电脑桌面采集同步合成方法步骤图。

具体实施方式

[0025] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作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。

[0026] 请参阅图1所示,本发明为一种视频直播图像声音采集和电脑桌面采集同步合成方法,包括如下步骤:

[0027] 步骤S1:获取在桌面预选的至少一个显示区域的音频/位置参数;

[0028] 步骤S2:抓取当前整个桌面视频信息进行绑定为第一数据包;

[0029] 步骤S3:摄像头采集教师上课视频进行绑定为第二数据包;

[0030] 步骤S4:将第一数据包和第二数据包进行解析合并;

[0031] 步骤S5:合并后的视频发送给服务器端进行区域音频/位置参数集的音频数据和视频图像进行匹配;

[0032] 步骤S6:匹配成功后的视频数据在指定区域位置进行覆盖。

[0033] 其中,步骤S1中,桌面视频和摄像头采集的上课视频可以在显示区域内进行切换。

[0034] 其中,步骤S2中,客户端截取整个桌面的视频图像,并进行编码压缩获取视频编码数据包;视频图像包括至少一个显示区域;视频编码数据包与至少一个位置参数绑定为对应的第二网络数据包,并发送给目标服务端。

[0035] 其中,步骤S4中,对桌面视频进行音频标志的标识处理,并形成区域音频/位置参数集,并将提取并编码压缩的音频数据与区域音频/位置参数集进行绑定为对应的第一网络数据包,并发送给目标服务端。

[0036] 其中,目标服务端对或去的第一网络数据包或第二网络数据包进行类型解析,若为第一网络数据包,则对编码压缩后的音频数据进行解码获取至少一组音频数据,若为第二网络数据包,则对编码压缩后的音频数据进行解码获取到视频图像。

[0037] 其中,目标服务端对位置参数对应的显示区域进行是否存在有音频标志的检测,若有,则根据音频标志对应的区域音频/位置参数集进行音频数据和位置参数对应的视频图像的匹配,并进行对应的客户端桌面区域共享,若无,则位置参数对应的视频图像进行对应的客户端桌面区域共享。

[0038] 其中,S5中,若第一数据包和第二数据包进行解析合并使均包含音频标志,则需要对合并后的音频数据进行混音处理。

[0039] 值得注意的是,上述系统实施例中,所包括的各个单元只是按照功能逻辑进行划分的,但并不局限于上述的划分,只要能够实现相应的功能即可;另外,各功能单元的具体名称也只是为了便于相互区分,并不用于限制本发明的保护范围。

[0040] 另外,本领域普通技术人员可以理解实现上述各实施例方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成,相应的程序可以存储于一计算机可读取存储介质中。

[0041] 以上公开的本发明优选实施例只是用于帮助阐述本发明。优选实施例并没有详尽叙述所有的细节,也不限制该发明仅为所述的具体实施方式。显然,根据本说明书的内容,可作很多的修改和变化。本说明书选取并具体描述这些实施例,是为了更好地解释本发明的原理和实际应用,从而使所属技术领域技术人员能很好地理解和利用本发明。本发明仅受权利要求书及其全部范围和等效物的限制。

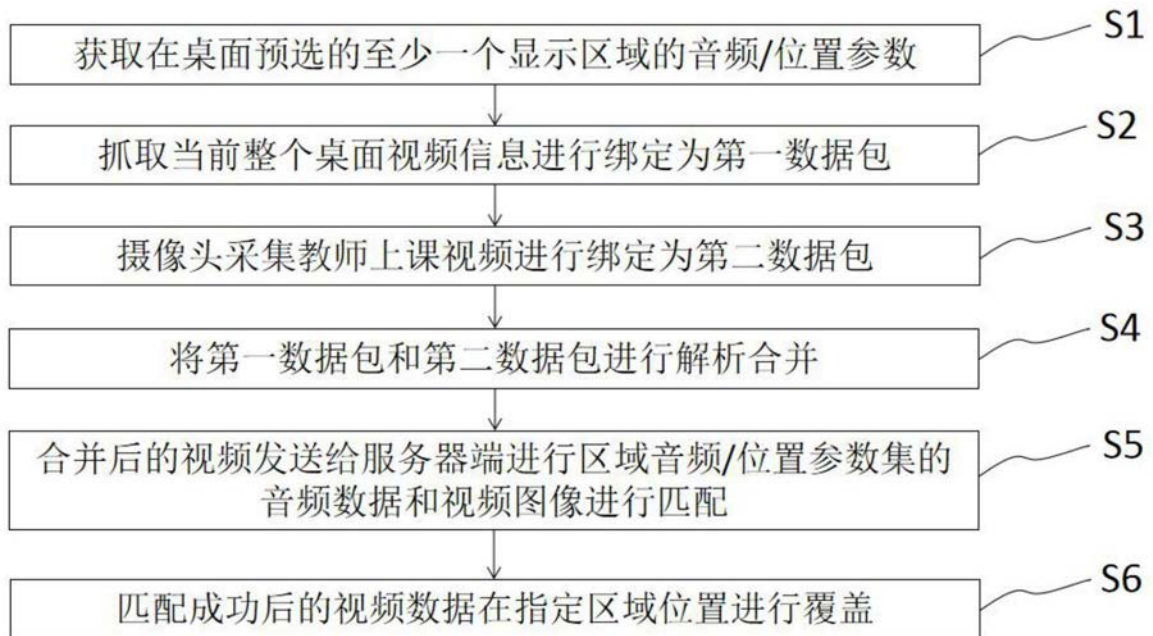


图1