



(12) 发明专利申请

(10) 申请公布号 CN 102084319 A

(43) 申请公布日 2011.06.01

(21) 申请号 200980125671.8

(22) 申请日 2009.05.05

(30) 优先权数据

61/058,459 2008.06.03 US

12/388,365 2009.02.18 US

(85) PCT申请进入国家阶段日

2010.12.31

(86) PCT申请的申请数据

PCT/US2009/042919 2009.05.05

(87) PCT申请的公布数据

W02009/148756 EN 2009.12.10

(71) 申请人 谷歌公司

地址 美国加利福尼亚州

(72) 发明人 M·芬克 R·琼 S·巴尔

A·巴尔齐赖 I·埃利亚斯

J·弗鲁马尔 H·霍 N·克雷姆

S·拉特纳 J·施罗克 R·塔沃里

I·科夫曼 A·布哈加瓦 V·王

L·维塞曼 S·穆图克里施南

M·巴多埃宇

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 王茂华

(51) Int. Cl.

G06F 3/00 (2006.01)

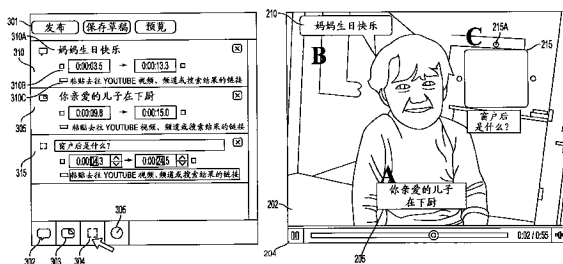
权利要求书 2 页 说明书 9 页 附图 5 页

(54) 发明名称

用于协作生成交互式视频的基于 WEB 的系统

(57) 摘要

提供了用于为已有在线托管的视频添加和显示交互式注释的系统和方法。图形注释界面允许创建注释以及将注释与视频相关联。注释可以是不同类型的,并且具有不同功能性,诸如改变已有视频的外观和/或行为,这例如是通过向已有视频补充文本、允许链接到其他视频或网页或者暂停视频的重放。对期望执行视频注释的用户的认证可以按照各种方式来执行,诸如通过对照已有列表检查统一资源定位符(URL),对照访问列表检查用户标识符,等等。作为认证的结果,给予用户适当的注释能力,诸如完整注释、不允许注释或者局限于视频的特定时间部分或空间部分的注释。



1. 一种用于对存储在视频储存库中的数字视频做注释的计算机实现的方法,所述方法包括:

确定是否授权没有向所述视频储存库贡献所述视频的用户对所述视频执行注释动作;

响应于所述用户被授权执行注释动作,向客户端提供用于对所述视频执行注释动作的用户界面;

接收对所述数字视频执行注释动作的请求;以及

执行所述注释动作,并且与所述数字视频相关联地存储所述注释动作的结果。

2. 根据权利要求1的计算机实现的方法,其中确定是否授权所述用户对视频执行注释动作包括:将与所述用户相关联的URL和与所述数字视频相关联的URL进行比较。

3. 根据权利要求1的计算机实现的方法,其中确定是否授权所述用户对视频执行注释动作包括:确定所述用户是否在被授权对所述视频执行注释的用户列表上。

4. 根据权利要求1的计算机实现的方法,其中确定是否授权所述用户对视频执行注释动作包括:监控所述数字视频的贡献者针对所述用户的动作。

5. 根据权利要求1的计算机实现的方法,其中确定是否授权所述用户对视频执行注释动作包括:监控其他用户针对所述用户的动作。

6. 根据权利要求1的计算机实现的方法,还包括:确定是否授权所述用户对所述视频的特定时间部分或空间部分执行注释动作。

7. 根据权利要求1的计算机实现的方法,其中确定是否授权所述用户对视频执行注释动作包括:确定所述视频的所有者是否已禁止对所述视频的注释。

8. 根据权利要求1的计算机实现的方法,其中所述注释动作包括创建注释,所述注释与所述视频的一个时间范围相关联并且具有可视外观。

9. 根据权利要求1的计算机实现的方法,其中所述注释动作包括创建针对所述视频内某个时刻的暂停注释,所述暂停注释使得当到达所述时刻时,所述视频的重放中断。

10. 根据权利要求1的计算机实现的方法,其中所述注释动作包括创建具有图形外观的注释,其中当所述注释在网络浏览器内被选择时,向不同的网页传递控制。

11. 根据权利要求1的计算机实现的方法,其中所述注释动作包括创建与所述视频的可视对象相关联的注释,所述注释具有图形外观,其中当所述注释被选择时,显示与所述可视对象相关联的信息。

12. 根据权利要求1的计算机实现的方法,还包括:

响应于从客户端接收针对所述视频的请求:

提供与所述视频相关联、并且能够由所述客户端显示的注释。

13. 根据权利要求1的计算机实现的方法,还包括:

响应于从客户端接收针对所述视频的请求:

改变将要在与创建注释的地区不同的地区显示的所述注释。

14. 根据权利要求1的计算机实现的方法,还包括:

响应于从阅览者接收针对所述视频的请求:

标识与所述视频相关联的注释;

通过阅览者的身份来过滤所标识的注释,由此产生经修改的注释集合;

向所述阅览者提供所述经修改的注释集合。

15. 根据权利要求 1 的计算机实现的方法,还包括:

响应于从阅览者接收针对所述视频的请求;

标识与所述视频相关联的注释;

通过所述阅览者所属群组的身位来过滤所标识的注释,由此产生经修改的注释集合;

向所述阅览者提供所述经修改的注释集合。

16. 根据权利要求 1 的计算机实现的方法,其中所述执行注释动作的请求包括根据光学字符识别得到的注释数据。

17. 根据权利要求 1 的计算机实现的方法,其中所述执行注释动作的请求包括根据图像对象识别得到的注释数据。

18. 一种存储有计算机程序的计算机可读存储介质,所述计算机程序可由处理器执行以对存储在视频储存库中的数字视频做注释,所述计算机程序的动作包括:

确定是否授权没有向所述视频储存库贡献所述视频的用户对所述视频执行注释动作;

响应于所述用户被授权执行注释动作,向客户端提供用于对所述视频执行注释动作的用户界面;

经由所提供的用户界面接收对所述数字视频执行注释动作的请求;以及

执行所述注释动作,并且与所述数字视频相关联地存储所述注释动作的结果。

19. 一种用于对存储在视频储存库中的数字视频做注释的计算机系统,所述计算机系统包括:

存储注释数据的注释数据库;以及

适合于执行动作的注释逻辑,所述动作包括:

确定是否授权没有向所述视频储存库贡献所述视频的用户对所述视频执行注释动作;

响应于所述用户被授权执行注释动作,向客户端提供用于对所述视频执行注释动作的用户界面;

经由所提供的用户界面接收对所述数字视频执行注释动作的请求;以及

执行所述注释动作,并且与所述数字视频相关联地将所述注释动作的结果存储在所述注释数据库中。

20. 一种用于对存储在视频储存库中的数字视频做注释的计算机实现的方法,所述方法包括:

接收利用针对所述数字视频的某个时刻的暂停注释对所述数字视频做注释的请求,所述暂停注释使得当到达所述时刻时,所述视频的重放中断;

创建所述暂停注释;以及

与所述数字视频以及数字视频的所述时刻相关联地存储所述暂停注释动作。

用于协作生成交互式视频的基于 WEB 的系统

技术领域

[0001] 所公开的实施方式总体上涉及用于协作生成数字视频的交互式特征。

背景技术

[0002] 传统的准许存储和显示数字视频的基于 web 的系统通常仅允许将视频作为整体进行评论。具体地,如果阅览者希望注释或以其他方式参考该视频的特定部分,则他们不得不通过视频中的文本或时间以及其他间接手段来显式地描述该部分。传统的系统还具有用于对视频做注释的简单控制,在这个意义上,它们完全允许这种注释。更确切地说,这种系统要么仅允许所有者(例如,上传视频的用户)添加注释,要么允许所有用户添加注释而没有限制。

发明内容

[0003] 本发明包括用于为在线托管的视频添加和显示交互式注释的系统和方法。图形注释界面允许创建注释以及将注释与视频相关联。注释可以是不同类型的,并且具有不同功能性,诸如改变已有视频的外观和/或行为,这例如是通过向已有视频补充文本、允许链接到其他视频或网页或者暂停视频的重放。

[0004] 对期望执行视频注释的用户的认证可以以各种方式来执行,诸如通过对照已有列表检查统一资源定位符(URL),对照访问列表检查用户标识符,等等。作为认证的结果,给予用户适当的注释能力,诸如完全注释、不允许注释或者局限于视频的特定时间部分或空间部分的注释。

[0005] 本发明内容部分以及下面的具体实施方式部分所描述的特征和优势并不是无所不包的。对于本领域普通技术人员而言,鉴于本文所提供的附图、说明书和权利要求,多种附加特征和优势将是很显然的。

附图说明

[0006] 图 1 是根据一个实施方式的用于允许对在线托管的视频做注释的系统架构的框图。

[0007] 图 2 示出了根据一个实施方式的可以添加到视频的不同类型的注释。

[0008] 图 3 绘出了根据一个实施方式的用于创建图 2 的注释的用户界面。

[0009] 图 4 示出了根据一个实施方式的向视频添加注释所涉及的步骤。

[0010] 图 5 示出了根据一个实施方式的允许添加注释以及提供有关已有注释的信息的注释界面。

[0011] 附图仅出于示意性目的而绘出本发明的各种实施方式。从下文讨论中本领域技术人员将很容易认识到,可以采用此处所示结构和方法的备选实施方式而不脱离在此描述的本发明的原理。

具体实施方式

[0012] 图 1 是根据一个实施方式的系统架构的框图。如图 1 所示,视频托管服务器 108 包括前端服务器 124、视频服务器 126、网络接口 122、视频数据库 128 和用户数据库 140。其他传统特征,诸如防火墙、负载均衡器、应用服务器、故障切换服务器、站点管理工具等并未示出,从而更清楚地图示出本系统的特征。用于实现系统的适当视频托管服务器 108 的示例包括 YouTube™ 网站和 Google Video™ 网站;已知还有其他视频托管站点,其可以适于按照此处公开的教导进行运作。应当理解,术语“网站”表示提供内容的任何系统和方法,并非旨在局限于支持经由因特网或 HTTP 协议提供内容的系统。各种服务器通常实现为单个软件或硬件,或者实现为多个软件或硬件,其可以经由网络接口 122 耦合到网络 105。一般而言,如果适合的话,在一个实施方式中描述为服务器侧所执行的功能在其他实施方式中也可以在客户端侧执行。

[0013] 客户端 130 运行浏览器 132 并经由网络 105 连接到前端服务器 124,网络 105 通常为因特网,但也可以是任何网络,包括但不限于 LAN、MAN、WAN、移动网络、有线或无线网络、专用网络或虚拟专用网络。尽管只示出了单个客户端 130 和浏览器 132,但是应当理解支持非常巨大数目(例如,数百万)的客户端,其可以在任何时间与视频托管服务器 108 通信。客户端 130 可以包括各种各样不同的计算设备。客户端设备 130 的示例有个人计算机、数字助理、个人数字助理、蜂窝电话、移动电话、智能电话或膝上型计算机。对于本领域普通技术人员而言显然的是,本发明不限于上面列出的设备。

[0014] 在有些实施方式中,浏览器 132 包括嵌入式视频播放器 134,诸如来自 Adobe Systems 公司的 Flash™ 播放器,或适于视频托管服务器 108 中使用的视频文件格式的任何其他播放器。用户可以通过浏览视频目录、进行关键词搜索、浏览其他用户或系统管理员的播放列表(例如,构成频道的视频的集合)或阅览与特定用户群(例如,社团)相关联的视频,来访问视频托管服务器 108 的视频。

[0015] 视频服务器 126 从内容提供者接收上传的媒体内容,并且允许客户端 130 阅览内容。可以从个人计算机经由因特网、从电话或 PDA 通过蜂窝网络,或者借助于本领域普通技术人员已知的用于通过网络 105 传送数据的其他手段将内容上传至视频服务器 126。内容可以按照类似的方式从视频服务器 126 进行下载;在一个实施方式中,将媒体内容提供作为向客户端 130 下载的文件;在一个备选实施方式中,将媒体内容流式传输到客户端 130。视频服务器 126 接收媒体内容所借助的手段不一定要与向客户端 130 递送内容所借助的手段匹配。例如,内容提供者可以经由个人计算机上的浏览器来上传视频,而客户端 130 可以阅览被作为流而发送到 PDA 的该视频。还应注意,视频服务器 126 本身可以用作内容提供者。客户端 130 与视频托管服务器 108 之间的通信,或者图 1 的其他不同单元之间的通信,可以是加密的或以其他方式编码的。

[0016] 客户端 130 的用户还可以基于关键词、标签或其他元数据来搜索视频。这些请求由前端服务器 124 作为查询来接收,并且被提供给视频服务器 126,视频服务器 126 负责在视频数据库 128 中搜索满足用户查询的视频。视频服务器 126 支持搜索视频的任何字段数据,包括题目、描述、标签、作者、类别等。

[0017] 客户端 130 和浏览器 132 的用户可以经由网络 105 向视频托管服务器 108 上传内容。上传的内容项例如可以包括视频、音频或视频与音频的组合。上传的内容被处理并被

存储在视频数据库 128 中。该处理可以包括格式转换（转码）、压缩、元数据标记以及其他数据处理。上传的内容文件与上传的用户相关联，因此将会根据需要更新用户数据库 140 中该用户的账户记录。

[0018] 出于一个实施方式的简便和描述的目的，上传的内容将称为“视频”、“视频文件”或“视频项”，但是此术语并不旨在对可以上传的内容的类型进行限制。每个上传的视频在其被处理时被指派视频标识符。

[0019] 用户数据库 140 负责维护浏览网站上的视频的所有用户的记录。每个个体用户被指派用户 ID（也称为用户身份）。用户 ID 可以基于任何标识信息，诸如用户的 IP 地址、用户名等。用户数据库还可以维护用户在视频上下文中以及通过其他应用（诸如使用电子邮件或文本消息收发）的声誉的有关信息。用户数据库可以进一步包含有关用户群中的成员资格的信息，用户群例如是可以浏览相同注释的用户的群组。用户数据库还可以包含针对给定用户的、被认为是该用户好友的其他用户的身份列表（此处使用的用于诸如授权用户列表、URL 列表等的术语“列表”广义地表示一组元素，其中各元素可以是有序的或无序的）。

[0020] 视频数据库 128 用于存储接收的视频。视频数据库 128 存储视频内容以及由其各自的内容所有者提供的相关联的元数据。视频文件具有与每个文件相关联的元数据，诸如视频 ID、艺术家、视频题目、标签、流派和时间长度。

[0021] 注释服务器 150 提供浏览视频数据库 128 中视频的注释以及向其添加注释的能力。注释服务器 150 根据用户或视频所有者的上传、发布者或者作为视频分析技术的结果来收集各种注释，诸如文本框、“思维气泡 (thought bubble)”等。注释服务器 150 继而将这些注释存储在注释数据库 154 中。对于给定视频，注释服务器 150 还向诸如客户端 130 或视频托管服务器 108 的实体提供注释数据库 154 中存储的针对该视频的注释。通常，注释修改在其他情况下是非交互式视频的行为，从而例如提供用户可与之交互的交互式覆盖 (overlay)，或者改变视频重放的常规流程。交互式覆盖的示例包括文本框、思维气泡、聚光灯、超链接、菜单、投票等，其中任何示例都可以具有任意复杂的用户界面行为。在一个实施方式中，注释服务器 150 位于与视频托管服务器 108 独立的物理服务器上，不过在其他实施方式中，注释功能可以包括在视频托管服务器 108 内。

[0022] 注释数据库 154 维护每个注释与所注释视频的适当部分之间的关联。在一个实施方式中，例如，注释数据库 154 存储注释类型（例如，文本框）的标识符以及与该类型相关联的任何信息（例如，文本标题）、应用该注释的视频的时间戳（例如，从时刻 01:05 到时刻 01:26）、该注释所注释的视频的标识符以及提交该注释的用户的标识符（例如，用户名）。某些类型的注释还可以与去往其他网页、视频、网络对象等的链接相关联。对于本领域技术人员而言，同样可以有用于注释的很多其他存储实现。

[0023] 视频分析模块 152 可以由注释服务器 150 用来自动地生成注释或者向用户建议注释。这可能需要诸如语音分析、视觉分析（例如，脸部检测、对象识别和光学字符识别 (OCR)）或者爬取 (crawl) 显式或隐式可用的注释等技术。

[0024] 由于视频注释可以由各用户通过网络 105 从远程位置来进行，因此可以使用认证机制将注释仅限制于用户子集。因此，提供认证服务器 170 以验证客户端 130 对注释服务器 150 的注释功能的访问权。如下面进一步描述的，在不同实施方式中，认证可以按照多种方式来执行，诸如使用机密链接、访问控制列表、用户信用得分或者基于社区限制的许可。在

一个实施方式中,采用三级许可系统,其中第一级是最低许可等级,用户仅可以通过点击链接来浏览视频注释以及与其交互,第二级是较高许可等级,用户可以添加或修改其自己的注释,第三级是最高许可等级,用户还可以修改和删除视频中的任何注释。机密链接的使用采用了 URL 列表 171,其将视频与 URL 相关联,通过该 URL 可以获得对注释界面的访问权。在一个实施方式中,认证服务器 170 实现为视频托管服务器 108 的部件。

[0025] 图 2 示出了根据一个实施方式的可以添加到视频的不同类型的交互式注释(此后称为“注释”)。客户端 130 上显示的主视频区 202 播放存储在视频数据库 128 中、由视频服务器 126 服务于客户端 130 的视频。可以经由例如视频控制区 204 来控制重放。在所示示例中,已经添加了三个不同的注释 205-215。注释 205 和 210 是文本框和思维气泡,其显示静态文本。注释 215 是聚光灯,其响应于用户将鼠标悬停在其边界之内而显示例如“窗户后是什么?”的文本。这些注释类型中的任何一个可以具有其活跃的时间范围,例如从时刻 0:15 到 0:36。例如,文本框 205 可以设置为在该视频播放 15 秒后出现,并且在用户有机会读它之后,于 21 秒后消失。

[0026] 这些注释类型中的任何一个还可以具有任意复杂的呈现,诸如形状和文本颜色和风格,或者关联的动作,诸如显示附加的注释或者在被激活后(例如通过鼠标点击、鼠标扫过、按压对应于该注释的键,等等),将用户重新定向到基于 web 的目标位置,诸如统一资源定位符(URL)。向其传递控制的目标位置可以包括广告,或者包括广告的内容。例如,点击聚光灯 215 可以去往描述特定产品的网页。目标位置还可以致使显示从不同角度获取的对象或场景,例如从不同照相机角度拍摄的对象的后侧。此外,目标位置可以具有链接、按钮或将控制传递回原始视频而不是其他视频的注释。在一个实施方式中,控制可以被传递回原始视频中的特定时刻,该时刻例如由对视频标识符和对视频中该时刻的描述进行编码的 URL 来指定,诸如“t = 0:22”,其表示进入视频后 22 秒的时间。时间戳在 URL 中的这种使用可以用于构建例如页面分支系列,其可以用于创建单个视频内的交互式故事情节。这例如允许快速转移到另一视频部分,而没有获取不同视频所带来的延迟。在一个实施方式中,可以有条件地显示注释,例如如果用户鼠标扫过另一注释,则在当时或者稍后显示该另一注释。

[0027] 还可以添加注释以修改视频的重放,而不是呈现交互式图形界面。例如,暂停注释使得视频的重放暂停一段给定时间延迟,包括无限长的时间。这例如允许用户在视频继续之前有任意的时间量来做出选择。使用如上所述的 URL 中的时间戳,可以修改视频的重放,从而例如点击门(或只是将鼠标放置在门上)将找到该视频中显示门打开以及其后的房间的那部分。这可以将视频中的互动级别增加到类似于计算机游戏的程度。

[0028] 对不同类型注释的使用可以用来以多种不同方式修改标准的线性视频浏览。注释可以用于实现例如菜单风格界面,其中视频通过具有去往其他页面的链接的注释来显示若干选项,并继而暂停该视频以允许用户选择一个选项。菜单项可以是静态注释、动画视频注释等等,并且可以作为单独标记的可视对象而在传统项目列表中显示,或者以各种其他方式显示。注释还可以用于实现分支故事情节,其中点击一个注释将通向该视频的一个续本,并且点击不同的注释通向不同的续本。例如,注释可以用于实现“剪刀、石头、布”的交互式游戏,其中例如点击对应于“石头”、“布”或“剪刀”选项的注释会通向分别描绘平局、赢或输的单独视频或同一视频的单独部分,每个结果潜在地导致显示代表该游戏的第二轮的附

加注释。菜单项还可以用于实现多角度故事情节,其中点击已做注释的角色脸部会导致从该角色的角度查看故事的剩余部分。

[0029] 图 3 绘出了根据一个实施方式的用于手动创建图 2 中注释的用户界面。注释图标 302-305 对应于四种注释类型(分别是语音气球、文本框、聚光灯以及暂停);选择其中之一并继而点击正在播放的视频,以便在对应于该点击的时间和位置处创建此类型的注释。注释继而具有默认值,诸如文本标题、时间范围、边界以及关联的 URL。在图 3 中,编辑区域 310、305 和 315 分别对应于显示的注释 205、210 和 215,并且其内容可以被编辑以改变标题的值。例如,编辑区域 310 包括文本标题 310A、时间范围 310B 和链接 310C。该链接例如可以是视频服务中的页面,其指示视频的观看页面,或者指示显示若干相关视频的缩略图的频道。文本标题 310A 已被用户设置为值“妈妈生日快乐 (Happy birthday mom)”,时间范围 310B 被设置为从 0:00:03 第 5 画面持续到 0:00:13 第三画面,而链接 310C 尚不具有指定值。对注释的编辑也可以通过图形方式完成;例如,标注 215 的边界可以通过拖拽与边界相关联的柄 (handle) 215A 来改变。

[0030] 作为使用图 3 的用户界面手动创建注释的备选或附加,图 1 的视频分析模块 152 可以用于自动地检测时间和空间位置以添加注释,确定其关联值,和 / 或控制已有注释的行为。

[0031] 这种分析的一个示例是脸部检测,其可以在图 3 的视频中使用,以将图像中男孩的脸部检测为人脸,并且建议创建与之关联的文本气球,或者其可以用于自动地提供或建议描述所识别脸部的标题。

[0032] 另一示例可以包括应用基于局部描述符匹配的对象识别方法。(例如,参见 Mikolajczyk, K., Schmid, C. 的“A Performance Evaluation of Local Descriptors”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 27, Issue 10 :1615-1630)。这种对象识别可以标识已知纹理的对象的实例,诸如位置、建筑物、书本、CD 封面等。(用于训练这种对象的识别的示例图像可以很容易地在产品 / 位置目录中找到,其将产品名与该产品的一个或多个图像相关联)。一旦检测到诸如特定 CD 的封面之类的对象,可以通过提供有关该对象的空间和时间定位的有根据的推测来简化手动注释过程。

[0033] 已识别的对象继而可以与注释相关联,所述注释诸如是从例如给定视点呈现更多信息的链接。例如,可以从国家角度呈现事件,其通过使用对象识别来标识与某个国籍相关联的对象并且呈现关联信息来实现,例如将某场板球赛的印度队队员与去往该队将参与的下一事件的链接相关联。作为另一示例,使用对象识别所识别的运动员可以与提供有关该运动员的统计数据、个人信息等的链接或其他注释数据相关联。

[0034] 此外,结合诸如 Google™ 或 YouTube™ 之类的搜索引擎的搜索索引,如果确实识别出视频中的对象,则可以参照该搜索索引以及作为针对该对象的链接而被建议的最前面的搜索结果,来执行描述该产品的短语(例如,搜索已识别音乐 CD 的题目,并且将其链接到对应于该题目的产品搜索页面)。

[0035] 在一个实施方式中,注释链接对应于搜索查询,因此如果用户点击该注释,则用户将看到该查询的搜索结果页面。例如,用户可以浏览视频中被用户标识的、其姓名已被用作搜索项的人员所发布的所有视频。此类注释允许结果页面保持最新,因为对与注释相关联的搜索项的搜索并非总是得到相同的结果页面。

[0036] 对象识别还可以用于标识感兴趣的位置,并且结合嵌入了位置信息的“带有地理标记的 (geotagged)”视频,可以将与识别出的位置有关的视频作为链接来提供或建议。

[0037] 通过跟踪对象在画面之间的运动,并由此将任何关联的注释与其一起移动,可以进一步增强对象识别。例如,如果男孩在视频各画面中移动其位置,则对象识别可以用于随着男孩的移动而跟踪男孩的脸部,并且自动地将文本气泡 210 重新定位在每个画面中检测到的脸部附近。这可以是一种类型的注释,其在画面内与画面中的对象一起移动。在Flash播放器的情况下,对视频的分析优选地可以在服务器上进行,而在视频播放期间不同画面位置中的注释的显示则通常随着对象在视频内移动而在播放器内实现。

[0038] 另一类型的分析是光学字符识别 (OCR),其细节对于本领域技术人员而言是公知的。例如,布告栏上的词语可以使用 OCR 来识别,并且为该布告栏自动地提供或建议对应的文本标题。

[0039] 图 4 示出了根据一个实施方式的向视频添加注释所涉及的步骤。客户端 130 使用网络 105 向视频服务器 108 请求视频。视频托管服务器 108 的前端服务器 124 接收该请求并授权给视频服务器 126,后者从视频数据库 128 获取视频并将其提供 410 给客户端 130。在一个实施方式中,视频托管服务器 108 继而授权注释服务器 150 提供注释用户界面;在其他实施方式中,视频托管服务器 108 向注释服务器 150 请求用于视频的注释,继而它自己提供注释用户界面,从而向注释服务器提供经由该用户界面创建的任何注释以供存储。在一个实施方式中,所提供的注释用户界面随做注释的用户的身份而不同。例如,如果做注释的用户不是视频的所有者(例如,向视频托管服务器 108 提交该视频的人员),则可以向该用户提供诸如已在图 3 中描述的编辑界面。在有些实施方式中,仅视频所有者或所有者指定的人员可以给视频做注释。在有些实施方式中,仅视频所有者或所有者指定的人员类别(例如,所有者的好友列表上的人员)可以给视频做注释。在有些实施方式中,任何人都可以给视频做注释。如果做注释的用户是视频的所有者,或者具有类似权限的其他用户,则可以提供诸如图 5 中所示的更全功能的注释界面。

[0040] 图 5 示出了既允许添加注释还提供与视频相关联的所有已有注释的有关信息的注释界面 500。在有些实施方式中,仅所有者可以使用这种界面。在其他实施方式中,更宽广的用户群可以使用这种界面。视频区 505 显示视频,注释编辑控件 507 允许以与图 3 中的控制 302-305 相类似的方式向视频添加注释。注释列表 510 显示当前与视频相关联的所有注释的列表,包括类型、关联的文本标题、向视频应用注释的时刻、提交注释的用户等等。可视时间线 515 图形化显示与该视频相关联的各个注释的时间位置。例如,标记 516 指示新添加的暂停注释的位置,并且拖动该标记 516 可以改变暂停的持续时间和/或时间位置。

[0041] 再次参考图 4,确定要提供哪种类型的注释用户界面(例如,所有者的、其他用户的)——其可以用来控制添加注释的程度——这可以包括使用认证服务器 170,其对用户进行认证 420。认证服务器在不同的实施方式中以不同的方式决定给定用户具有哪种注释访问权。在一个实施方式中,认证服务器 170 仅允许由视频的所有者创建注释。在此实施方式中,认证服务器 170 查询用户数据库 140 以确定该视频所有者的身份(例如,用户名),并且将其与当前活跃用户的身份进行比较。如果当前活跃用户没有身份(例如,没有登记),或者身份与视频所有者的身份不匹配,则不提供注释能力。

[0042] 在另一实施方式中,视频所有者通过向用户提供与该视频相关联的专用 URL 来授

予该用户访问权。例如,再次参考图 5,所有者界面 500 包括 URL 区 520,其列出了 URL 列表 171 中与该视频相关联的 URL。所有者继而可以复制此 URL 并根据需要分发给该所有者希望其具有注释能力的其他用户,例如通过电子邮件将该 URL 发送给好友,将其贴在朋友可访问的网页上,等等。接着,当用户进入该 URL,认证服务器 170 将该 URL 与 URL 列表 171 中用于给定视频的 URL 进行比较;如果 URL 匹配,则提供注释用户界面。在 URL 变得分布范围过大并且过多不希望的注释被添加到视频上时,所有者可以使用按钮 525 来禁止更多注释的添加,在这种情况下,URL 变为不可用于进一步注释的目的。

[0043] 在又一实施方式中,所有者可以向视频托管服务器 108 的特定用户授予注释许可,例如通过指定授权用户名列表。许可可以是可做注释,在这种情况下,列表上的任何用户在浏览该视频时将被提供以注释界面。在另一实施方式中,许可是建议注释,在这种情况下,用户界面将允许添加尝试性注释但在重放期间不实际显示,直到所有者或其他授权用户同意为止。许可还可以限于允许对视频的特定时间部分或物理区的注释。这可以这样实现:允许所有者使用由注释编辑器提供的相同空间和时间控件在其视频中定义“协作区域”,例如,通过使用诸如图 3 的时间范围控件 310B 之类的控件来指定时间范围,以及通过以图形方式在视频区 202 中选择区域来指定一个区。定义协作区域的数据继而可以存储在例如注释数据库 154 中。利用所定义的协作区域,允许被确定为视频所有者的协作者的其他用户(例如,被认为是该用户的“好友”的用户)向视频添加注释,但是仅在指定的协作区域的空间和时间范围内。

[0044] 在又一实施方式中,访问控制列表无需由所有者或其他授权用户严格地预先定义,而是可以基于外部因素而动态地确定。例如,列表可以基于用户的信用得分(信用得分可以例如通过监控此贡献者有多少在先注释被内容所有者删除来计算,或者通过社区限制动作来计算,这些动作诸如用户表达同意或不同意用户的视频评级,诸如通过“赞同”或“否决”指示符),或者基于用户数据 140 中被认为是该用户的好友的那些用户(例如,基于诸如活跃电子邮件交换的行为,或者基于友谊的明确指示,诸如将某人添加到“好友”列表)的信用得分。

[0045] 在另一实施方式中,以分层方法定义注释,其中注释的创建者确定注释在视频重放时刻的可见性。在此方法中,不同的用户或用户群可以自由地定义其自己的注释,但是得到的注释在播放视频时将仅对那些用户或用户群是可见的。

[0046] 在认证服务器 170 已确定授权用户查看哪种注释界面之后(诸如图 3 或图 5 的界面,或者不允许创建注释的基本重放界面),注释服务器 150 继而向用户提供 430 该注释界面,可选地还一起提供该视频的任何预先存在的注释。如果界面允许创建注释,则用户使用该界面来指定注释的属性。结果继而例如响应于用户在指定期望的注释之后选择图 5 的发布按钮 525 而被传送至视频数据库 128,视频数据库 128 随后接收 440 并存储这些注释。

[0047] 当浏览已注释的视频时,发生类似的过程。首先,客户端 130 使用网络 105 向视频服务器 108 请求视频。视频托管服务器 108 的前端服务器 124 接收该请求并授权给视频服务器 126,后者从视频数据库 128 获取视频。视频托管服务器 108 继而从注释服务器的注释数据库 154 获取适合的注释,可选地,这受制于认证服务器 170 对用户的认证。在一个实施方式中,视频托管服务器随着视频的加载而获取用于该视频的所有注释。在其他实施方式中,注释随着视频的前进而流式传输,或者在若干预定义的块(chunk)中发送。

[0048] 适合的注释可以在不同实施方式中有所变化。在一个实施方式中,提供所有注释而不考虑用户的身份。在采用层的另一实施方式中,仅提供浏览该视频的用户或用户所属群组的成员所创建的注释。注释服务器 150 还可以针对特定客户端 130、基于有关该客户端的知识将注释转换为适当的格式;例如,如果客户端处于语言不同于注释的语言的地区,则可以执行对注释文本的翻译,而且可以禁止客户端不支持的注释,诸如对于限于文本的简单客户端禁止动画视频注释。

[0049] 利用已选择以及适当格式化的适当注释集合,注释服务器 150 继而向客户端 130 提供这些注释,客户端 130 结合播放的视频显示这些注释,从而修改视频的外观和/或行为,例如使用上面结合图 2 和图 3 描述的注释类型。例如,如果客户端 130 遇到图 2 中的任何注释类型 205、210、215,则其将文本标签与视频中该注释所指定的部分关联显示。如果注释具有关联的超链接,则客户端 130 将显示该超链接并将采取动作,诸如在网络浏览器(例如,当前呈现该视频和注释的网络浏览器窗口)内显示通过该超链接所链接至的网页。如果注释是“暂停”类型,并且客户端 130 到达该暂停注释所对应的时间周期的起点,则客户端 130 在该暂停注释所指定的时间段内中断视频的重放,接着在该指定的时间段到期后继续重放。

[0050] 应当理解,上述描述中确切的部件以及布置、操作顺序以及其他方法纯粹是出于示例目的,本领域技术人员同样可以想到更多各种各样的备选部件布置及操作顺序。例如,图 1 的注释服务器 150 可以是视频托管服务器 108 的一部分,与视频托管服务器位于同一本地网络上的服务器,或者位于视频托管服务器的远程网络上。作为另一示例,认证服务器 170 可以与视频托管服务器分离,或者可以是其一个部件。此外,有些客户端 130 可以配置为不与注释服务器 150 通信,或者注释服务器 150 不向客户端 130 提供注释,这种情况下,客户端 130 获取未注释的视频(例如,各个国家或各种语言的客户端可以具有其自己的注释设置)。在有些实施方式中,客户端 130 可以离线进行注释,而无需网络连接到注释服务器 150,之后在网络连接可用时将这些注释与注释服务器 150 同步。

[0051] 本说明书中所提及的“一个实施方式”或“实施方式”是指包含在本发明的至少一个实施方式与实施方式相结合描述的特定特征、结构或特性。在说明书中各处出现的短语“在一个实施方式中”不一定所有都指代同一实施方式。

[0052] 应当注意,本发明的过程步骤和指令可以具体化在软件、固件或硬件中,当其具体化在软件中时,可以被下载以驻留在由各种操作系统所使用的不同平台上或从其进行操作。

[0053] 本发明还涉及用于执行此处的操作的装置。该装置可以是专为所需目的而构建的,或者其可以包括由存储在计算机中的计算机程序来选择性激活或重新配置的通用计算机。这种计算机程序可以存储在计算机可读存储介质中,诸如但不限于,包括软盘、光盘、CD-ROM、光磁盘的任意类型的盘、只读存储器 (ROM)、随机访问存储器 (RAM)、EPROM、EEPROM、磁卡或光卡、专用集成电路 (ASIC) 或适于存储电子指令的任意类型的介质,并且每一个都耦合至计算机系统总线。而且,本说明书中所涉及的计算机可以包括单处理器或者可以是采用多处理器设计以增强计算能力的架构。

[0054] 本文所提供的算法和显示本质上并不涉及任何特定计算机或其他装置。也可以使用具有按照本文教导的程序的程序的各种通用系统,或者构建更为专用的装置来执行所需方法步

骤可能更为方便。以下描述中将呈现用于各种各样的这些系统的所需结构。此外,本发明并未参考任何特定编程语言进行描述。将会理解,各种各样的编程语言可以用于实施本文所描述的本发明的教导,并且提供以下对任何特定语言的参考以用于本发明的可实施性和最佳模式的公开。

[0055] 尽管已经参考优选实施方式和若干备选实施方式特别描述和示出了本发明,但是相关领域技术人员将会理解,可以在其中做出各种形式和细节上的改变而不偏离本发明的精神和范围。

[0056] 最后,应当注意,本说明书中所使用的语言主要是选择以用于易读和指导目的,而不是选择以描绘或约束该创造性主题。相应地,本发明的公开旨在于示意性目的,而不是限制本发明的范围,本发明的范围由所附权利要求来阐述。

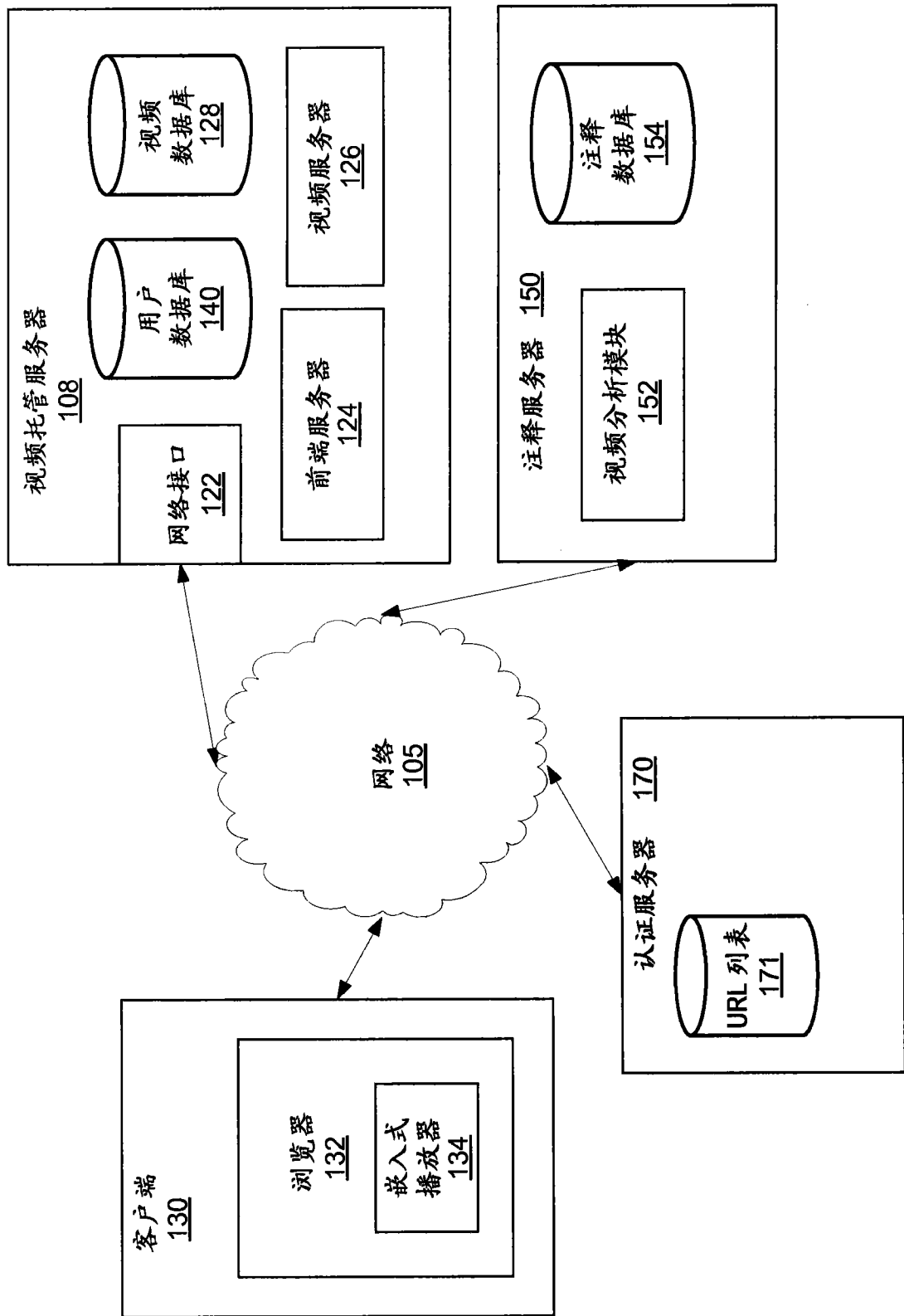


图 1

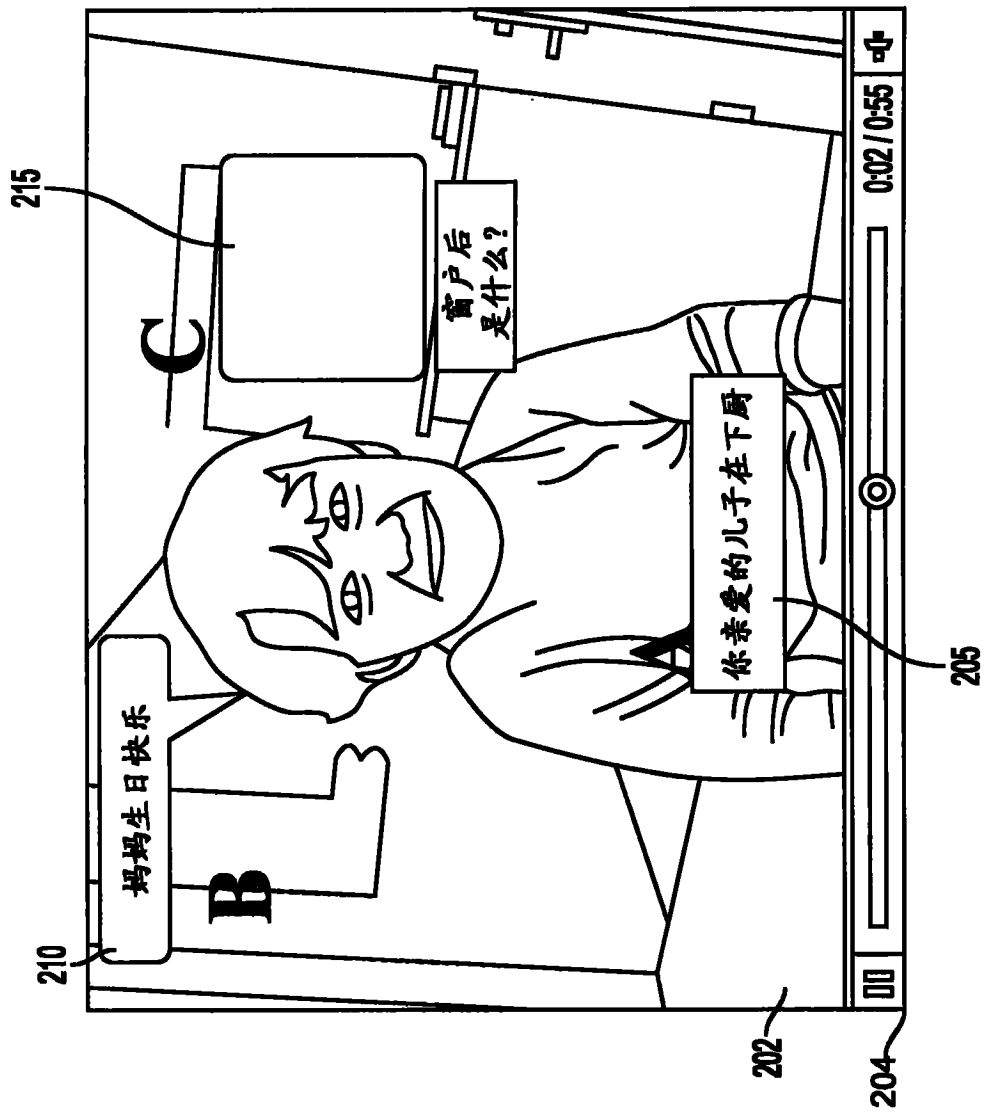


图 2

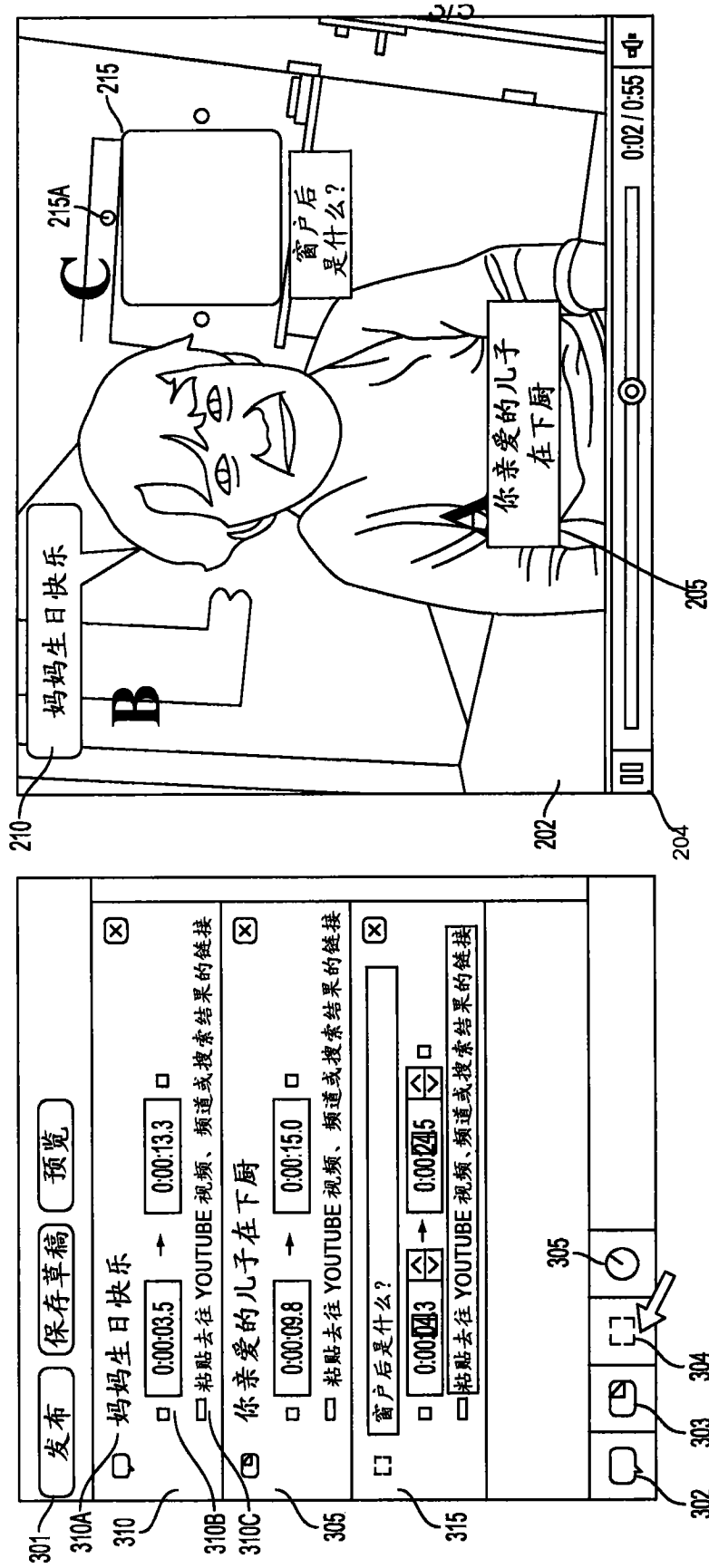


图 3

400

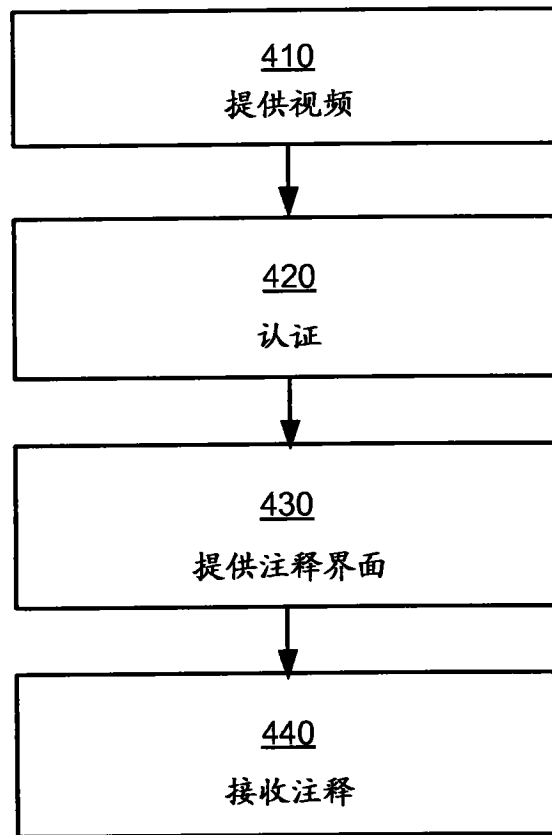


图 4

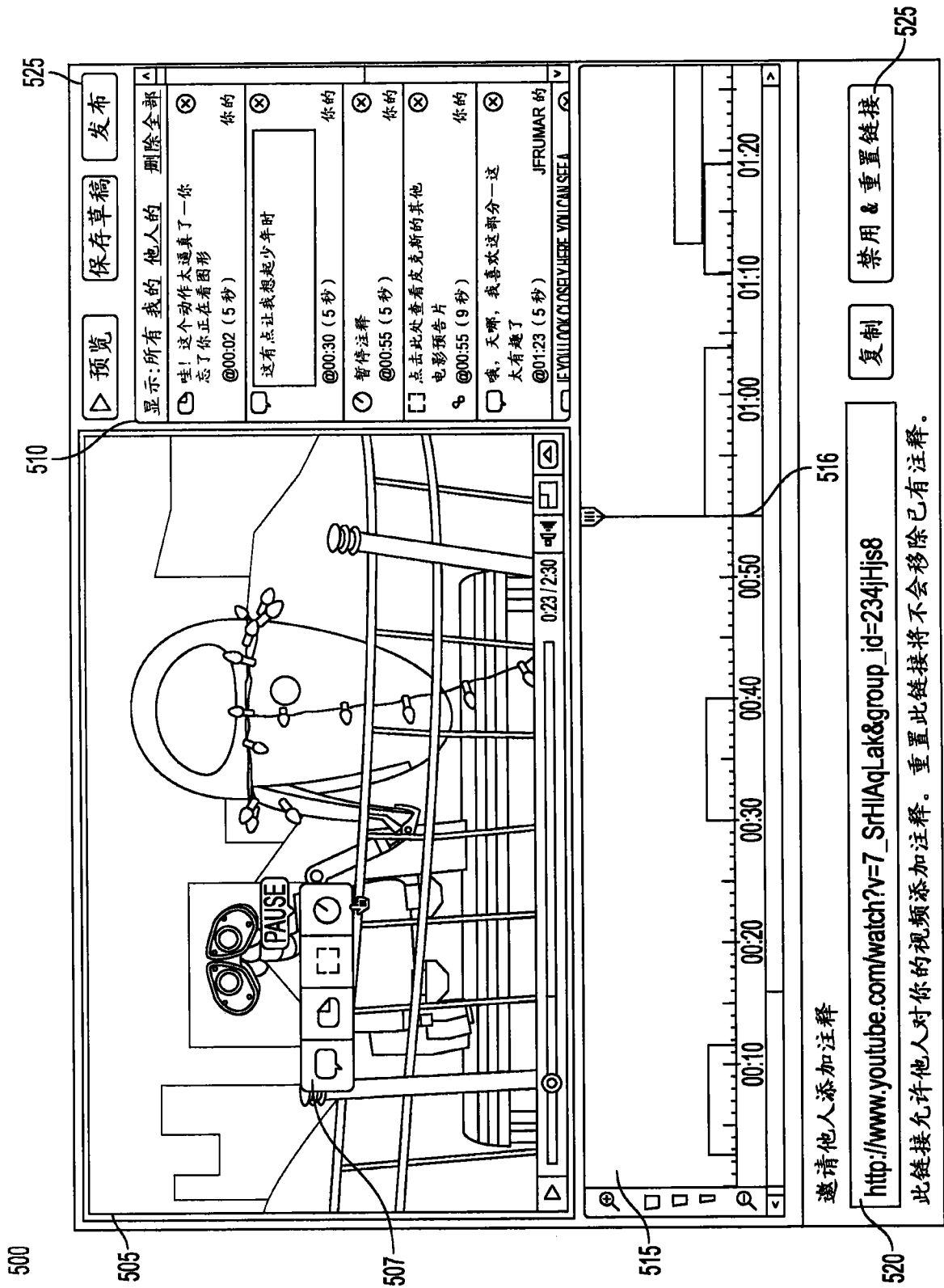


图 5