



(12)发明专利

(10)授权公告号 CN 105474213 B

(45)授权公告日 2019.12.03

(21)申请号 201480046451.7

(22)申请日 2014.07.28

(65)同一申请的已公布的文献号
申请公布号 CN 105474213 A

(43)申请公布日 2016.04.06

(30)优先权数据
61/859,941 2013.07.30 US
61/859,935 2013.07.30 US

(85)PCT国际申请进入国家阶段日
2016.02.22

(86)PCT国际申请的申请数据
PCT/US2014/048355 2014.07.28

(87)PCT国际申请的公布数据
W02015/017304 EN 2015.02.05

(73)专利权人 柯达阿拉里斯股份有限公司
地址 美国纽约州

(72)发明人 A·C·路易 J·A·马尼科

(74)专利代理机构 北京律盟知识产权代理有限
责任公司 11287
代理人 齐杨

(51)Int.Cl.
G06F 16/583(2019.01)
G06K 9/00(2006.01)
G06K 9/46(2006.01)

(56)对比文件
WO 2009155991 A1,2009.12.30,
CN 103186917 A,2013.07.03,

审查员 袁冠群

权利要求书3页 说明书7页 附图9页

(54)发明名称

用于创建可操纵视图的系统及方法

(57)摘要

本发明涉及一种用于创建可操纵视图的方法,其包含:接收数字图像;计算所述数字图像中的每一者的一组特征点;选择所述数字图像中的一者作为参考图像;识别所述参考图像中显著的所关注的区域;使用针对其它数字图像中的每一者而计算的所述组特征点来识别含有与所述参考图像中所述显著的所关注的区域类似的所关注的区域的其它数字图像;指定针对所述参考图像中所述显著的所关注的区域的参考位置;将所述其它数字图像与含有所述经指定的参考位置的所述图像对准;将含有所述经指定的参考位置的所述图像及所述其它数字图像进行排序;以及产生可操纵视图。



1. 一种用于使用多个经排序的图像来创建可操纵视图的方法,其包括:

接收多个数字图像,所述多个数字图像中的每一者包含人脸描绘及至少一个不同于所述人脸描绘的静止背景对象;

使用处理器来计算所述多个数字图像中的每一者的一组特征点;

选择所述多个数字图像中的一者作为参考图像;

识别所述参考图像中显著的所关注的区域;

使用所述处理器来使用针对所述多个数字图像而计算的所述一组特征点来识别所述多个数字图像中的一组数字图像,所述一组数字图像中的每一者包含与所述参考图像中所述显著的所关注的区域相同的所关注的区域;

指定针对所述参考图像或所述一组多个数字图像中的一者中的所述显著的所关注的区域的参考位置;

使用所述处理器来将所述参考图像及所述一组多个数字图像中的每一者相对于经指定的参考位置对准;

至少部分地基于所述多个数字图像的捕获时间将所述参考图像及所述一组多个数字图像中的每一者进行排序;以及

产生包括显示所述参考图像及执行从所述参考图像到所述一组多个数字图像中的一者幻灯片过渡的可操纵视图,

其中所述显著的所关注的区域包括所述参考图像中的所述至少一个静止背景对象。

2. 根据权利要求1所述的方法,其中所述特征点包括尺度不变特征变换SIFT关键点、加速稳健特征SURF关键点、定向梯度直方图HOG描述符、梯度位置及定向直方图GLOH描述符。

3. 根据权利要求1所述的方法,其中选择所述多个数字图像中的一者作为所述参考图像包括接收用户输入作为所述参考图像的选择。

4. 根据权利要求1所述的方法,其中选择所述多个数字图像中的一者作为所述参考图像进一步包含:

使用所述处理器基于所述经计算的一组特征点来识别所述多个数字图像中的一组类似图像;以及

选择所述经识别的一组类似图像中的一者作为所述参考图像。

5. 根据权利要求1所述的方法,其中识别所述参考图像中所述显著的所关注的区域包括使用所述处理器基于所述参考图像的所述经计算的一组特征点来识别所述显著的所关注的区域。

6. 根据权利要求1所述的方法,其中识别所述参考图像中所述显著的所关注的区域包括接收用户输入作为所述显著的所关注的区域的指示。

7. 根据权利要求1所述的方法,其中对准包含使用旋转、移位、按比例调整、缩放、裁剪、透视变换或其组合来将所述参考图像与所述一组多个数字图像中的每一者几何对准。

8. 根据权利要求1所述的方法,其中对准包含使用色彩平衡、色阶调整、亮度增强、对比度增强、噪声调整或其组合来将所述参考图像与所述一组多个数字图像中的每一者光度对准。

9. 根据权利要求1所述的方法,其中对准进一步包含使用所述处理器将对应于特殊效应的滤波器应用到所述一组多个数字图像或所述参考图像。

10. 根据权利要求1所述的方法,其中所述参考图像及所述一组多个数字图像按时间顺序、所述多个数字图像中的每一者中的人物的数目、所述多个数字图像中的每一者中的每一人物的年龄、所述多个数字图像中的每一者中的人物的身份或其组合而排序。

11. 根据权利要求1所述的方法,其中所述可操纵视图包括幻灯片演示、经排序的参考图像及所述一组多个数字图像的视频、所述经排序的参考图像及所述一组多个数字图像的全景图、交互式演示、动画演示、交互式游戏或虚拟现实环境。

12. 一种用于使用多个经排序的图像来创建可操纵视图的系统,其包括:

存储器,其经配置以存储多个数字图像,所述多个数字图像中的每一者包含人脸描绘及至少一个不同于所述人脸描绘的静止背景对象;以及

处理器,其经配置以:

计算所述多个数字图像中的每一者的一组特征点;

接收对所述多个数字图像中的一者作为参考图像的选择;

接收所述参考图像中显著的所关注的区域的识别;

使用针对所述多个数字图像而计算的所述一组特征点来识别所述多个数字图像中的一组数字图像,所述一组数字图像中的每一者包含与所述参考图像中所述显著的所关注的区域相同的所关注的区域;

指定针对所述一组多个数字图像中的一者中所述显著的所关注的区域的参考位置;

将所述参考图像及所述一组多个数字图像中的每一者相对于经指定的参考位置对准;

至少部分地基于所述多个数字图像的捕获时间将所述参考图像及所述一组多个数字图像中的每一者进行排序;

产生包括显示所述参考图像及执行从所述参考图像到所述一组多个数字图像中的一者幻灯片过渡的可操纵视图;以及

将经产生的可操纵视图存储至所述存储器;

其中所述显著的所关注的区域包括所述参考图像中的所述至少一个静止背景对象。

13. 根据权利要求12所述的系统,其中所述特征点包括尺度不变特征变换SIFT关键点、加速稳健特征SURF关键点、定向梯度直方图HOG描述符、梯度位置及定向直方图GLOH描述符。

14. 根据权利要求12所述的系统,其中所述处理器经配置以接收用户输入作为所述参考图像的选择。

15. 根据权利要求12所述的系统,其中所述处理器进一步经配置以:

基于所述经计算的一组特征点来识别所述多个数字图像中的一组类似图像;以及

选择所述一组类似图像中的一者作为所述参考图像。

16. 根据权利要求12所述的系统,其中基于所述参考图像的所述经计算的一组特征点来识别所述参考图像中所述显著的所关注的区域。

17. 根据权利要求12所述的系统,其中所述处理器经配置以接收用户输入作为所述参考图像中所述显著的所关注的区域的指示。

18. 根据权利要求12所述的系统,其中所述处理器经配置以使用旋转、移位、按比例调整、缩放、裁剪、透视变换或其组合来几何对准所述参考图像及所述一组多个数字图像中的每一者。

19. 根据权利要求12所述的系统,其中使用色彩平衡、色阶调整、亮度增强、对比度增强、噪声调整或其组合将所述参考图像及所述一组多个数字参考图像中的每一者光度对准。

20. 根据权利要求12所述的系统,其中所述处理器经配置以将对应于特殊效应的滤波器应用到所述一组多个数字图像或所述参考图像。

21. 根据权利要求12所述的系统,其中所述参考图像及所述一组多个数字图像按时间顺序、所述多个数字图像中的每一者中的人物的数目、所述多个数字图像中的每一者中的每一人物的年龄、所述多个数字图像中的每一者中的人物的身份或其组合而排序。

22. 根据权利要求12所述的系统,其中所述可操纵视图包括幻灯片演示、经排序的参考图像及所述一组多个数字图像的视频、所述经排序的参考图像及所述一组多个数字图像的全景图、交互式演示、动画演示、交互式游戏或虚拟现实环境。

用于创建可操纵视图的系统及方法

[0001] 相关申请案的交叉参考

[0002] 本申请案涉及2013年7月30日申请的标题为“用于创建可操纵视图的方法 (METHOD FOR CREATING NAVIGABLE VIEWS)”序列号为61/859,935的美国临时专利申请案及2013年7月30日申请的标题为“用于创建可操纵视图的系统 (SYSTEM FOR CREATING NAVIGABLE VIEWS)”序列号为61/859,941的美国临时专利申请案,且本申请案主张前述两个临时专利申请案的优先权。上述共同待决的临时申请案的揭示内容特此以全文引用的方式并入。

技术领域

[0003] 本发明大体上涉及数字图像处理领域,且特定来说,涉及用于创建消费者媒体集中的可操纵视图及演示的系统及方法。

背景技术

[0004] 数码相机及智能电话的激增已导致大量数字图像及视频的出现,从而产生大的个人图像及视频数据库。由于拍摄数字照片及视频是容易的且几乎免费的,因此消费者不再将拍摄照片限于重要事件及特殊场合。图像及视频被频繁地捕获,且成为消费者日常生活中每天都要发生的事情。由于一般用户已积累许多年的数字图像及视频,因此浏览数字图像及视频的集合来创建幻灯片放映或可操纵多媒体演示对于消费者来说是耗时的过程。

[0005] 从消费者图像及包含音乐声轨的视频帧汇编而成的数字幻灯片放映是众所周知的。与从模拟到数字摄像的许多转换形式一样,数字幻灯片未能相对于其模拟对应物在很大程度上增强用户体验,且未充分利用数字处理及媒体的潜能。除了数字转换及音乐节拍导出的演示时序之外,图像随着声轨播放而按顺序呈现,与模拟幻灯片放映一样。用户通常拍摄人物及对象的照片,而将背景作为次要考虑因素,除非用户在拍摄日落或风景。由于人物及对象提供摄像的主要动机,因此其还应作为呈现所得图像的主要考虑因素。图像编辑软件(例如,Adobe“后效应”(Adobe“After Effects”))用于创建运动图形及视觉效应且允许用户借助于各种内置工具及第三方插件以及对如视差的变量的个人关注及用户可调整的观察角度而在2D及3D空间中把媒体制成动画效果、改变及合成媒体。问题在于学习及使用如上述图像编辑软件的工具非常复杂,且需要熟练的专业图形绘制技术人员。使用经提取的对象、人物或所关注的区域、基于主题的“脚本”、及经导引的用户选择、多个选择选项,创建动画的/可操纵的多媒体演示的整个过程可为自动的。本发明提供一种自动化方法以使创建针对用户的图像集合的有趣且个性化的多媒体演示及可操纵视图的繁琐过程更容易。

发明内容

[0006] 根据本发明,提供一种用于使用多个经排序的图像来创建可操纵视图的系统及方法,其包括:接收多个数字图像;使用处理器来计算所述多个数字图像中的每一者的一组特征点;选择所述多个数字图像中的一者作为参考图像;识别所述参考图像中显著的所关注

的区域;使用所述处理器来使用针对其它数字图像中的每一者而计算的所述组特征点来识别含有与所述参考图像中所述显著的所关注的区域类似的所关注的区域的所述多个图像中的其它数字图像;指定针对所述参考图像或所述经识别的一组类似图像中的一者中所述显著的所关注的区域的参考位置;使用所述处理器来将所述其它数字图像与含有所述经指定的参考位置的图像对准;将含有所述经指定的参考位置的图像及所述其它数字图像进行排序;以及产生显示所述经排序的数字图像的可操纵视图。

[0007] 现有消费者图像集合可经组织及集成以创建虚拟环境。通过使用计算机视觉及模式辨识技术,图像集合的演示及操纵变成与常规“模拟式”数字幻灯片放映相反的交互式、沉浸式、非线性体验。本发明旨在通过对数字图像的群组中显著的反复出现的所关注的区域(ROI)或对象的进行识别来促进可操纵多媒体演示。通过将参考图像中的关键点与目标图像中的关键点进行匹配来识别显著的反复出现的ROI或对象。因此,冗余、较不受关注的图像可获得作为用于创建虚拟背景的潜在材料的价值。这与现有图像集合(即,无需准备或特殊捕获需求)兼容。此外,反复出现的静止对象可用作到额外信息或操纵点的超链接。用户可独立地选择背景及对象以创建适于打印及共享的场景。

附图说明

- [0008] 图1展示根据本发明的实施例的系统的框图;
- [0009] 图2描述根据本发明的实施例的方法的流程图;
- [0010] 图3描述图2的框102的流程图;
- [0011] 图4展示根据本发明的实施例的出现在两个不同图像中的同一对象的SIFT关键点的匹配的实例;
- [0012] 图5A展示根据本发明的实施例的按时间顺序的图像顺序;
- [0013] 图5B展示根据本发明的实施例的按人物的数目来排序的图像;
- [0014] 图5C展示根据本发明的实施例的按年龄来排序的具有单个人物的图像;
- [0015] 图5D展示根据本发明的实施例的按时间来排序的具有同一人物的图像;及
- [0016] 图6展示根据本发明的实施例的经由缩放及裁剪的对准的实例。

具体实施方式

[0017] 如所属领域的技术人员将熟知,本发明可在计算机系统中实施。在以下描述中,本发明的一些实施例将被描述为软件程序。所属领域的技术人员将容易认识到,此方法的等效物还可被解释为本发明范围内的硬件或软件。

[0018] 因为图像处理算法及系统是众所周知的,所以本描述将尤其涉及形成根据本发明的方法的部分或更直接地与根据本发明的方法协作的算法及系统。此类算法及系统的其它方面,以及用于产生及以其它方式处理其中所涉及的图像信号的硬件或软件(本文未特定展示或描述)可从所属领域已知的此类系统、算法、组件及元件中选择。鉴于如以下说明书中所阐述的描述,本发明的所有软件实施方案为常规的且在所属领域的技术人员理解范围之内。

[0019] 集合中的视频在描述的其余部分中包含于术语“图像”中。

[0020] 本发明可在计算机硬件及计算机化设备中实施。举例来说,可在数码相机、多媒体

智能电话、数字打印机中执行所述方法,可在因特网服务器、自助服务终端及个人计算机上执行所述方法。参考图1,图1说明用于实施本发明的计算机系统。尽管出于说明根据本发明的实施例的目的展示所述计算机系统,但本发明不限于所展示的计算机系统,而是可用在任何电子处理系统(例如,数码相机、家用计算机、自助服务终端中的电子处理系统)或用于数字图像处理的任何其它系统上。计算机10包含基于微处理器的单元20(本文也称作处理器),其用于接收及处理软件程序且用于执行其它处理功能。存储器单元30存储用户供应及计算机生成的数据,所述数据可由处理器20在运行计算机程序时存取。显示装置(例如,监视器)70(例如,通过图形用户接口)电连接到计算机10用于显示与软件相关联的信息及数据。键盘60也连接到所述计算机。作为对使用键盘60进行输入的替代,可使用鼠标来移动显示装置70上的选择器且用于选择所述选择器覆盖的项,如在所属领域中为众所周知的。输入装置50(例如,光盘(CD)及DVD)可插入计算机10以用于将软件程序及其它信息输入计算机10及处理器20。此外,如在所属领域中为众所周知的,计算机10可经编程以用于内部存储软件程序。此外,媒体文件(例如,图像、音乐及视频)可通过输入装置50(例如,存储卡、随身盘、CD及DVD)被传送到计算机10的存储器30,或通过将捕获装置(例如,摄像机、蜂窝电话及视频记录器)作为输入装置直接连接到计算机10来将所述媒体文件传送到计算机10的存储器30。计算机10可具有到外部网络(例如,局域网或因特网)的网络连接(例如,电话线或无线连接80)。软件程序及媒体文件可通过所述网络连接从其它计算机或因特网传送到所述计算机。

[0021] 还应注意,本发明可在软件或硬件的组合中实施且不限于物理连接或定位在同一物理位置内的装置。图1中所说明的装置中的一或多者可远程定位且可经由网络连接。所述装置中的一或多者可(例如)通过射频链路直接无线连接或经由网络无线连接。

[0022] 参考图2,在步骤100中接收一组数字图像。通过将在所述组数字图像中的参考图像中的关键点与目标图像中的关键点进行匹配来识别反复出现的显著的ROI或对象。通过首先计算图像组中的每一图像的特征点(例如,SIFT关键点)而在步骤101中实现此种情况。随后在步骤102中,在从数字图像组选择的参考图像中识别反复出现的显著的ROI或对象。在步骤103中,应用景象匹配算法以识别具有如较早识别的类似显著的ROI或对象的所有其它图像。接下来,指定在经识别的一组类似图像中的一者中显著的ROI或对象的参考位置104。接着,相对于显著的ROI或对象的参考位置而对准(例如,旋转、移位或裁剪)经识别的图像中的每一者的对应ROI或对象105。在所述对准之后,将所述组类似图像以时间方式(或以其它方式,例如按人物的数目或人物的年龄)进行排序106。最后,使用经排序的一组图像来创建演示或可操纵视图107。以下描述各个步骤的细节。

[0023] 从所述图像提取的优选特征为SIFT特征,由劳(Lowe)在国际计算机视觉期刊(International Journal of Computer Vision),2004年第2期,60卷中描述。此方法使用基于高斯差(difference-of-Gaussians)的最大值及最小值的关键点选择,且使用区域101中的梯度量值及定向来描述这些关键点。所述关键点通常表示图像中所关注的边缘及拐角。劳(Lowe)还描述用于使用SIFT来匹配模式(即使在所述图像中的其它部分改变且所述模式的尺度及定向发生改变时)的方法。也可匹配图像中的所关注的点(其中尺度可不同)的对SIFT的替代是由贝(Bay)等人在2006年的第九次欧洲计算机视觉会议(ECCV)中提出的SURF(加速稳健特征)。这是基于有效使用积分图像的2D Haar小波响应的总和的更快方法。

其它特征可包含HOG(定向梯度直方图)描述符及GLOH(梯度位置及定向直方图)描述符。由克里斯蒂安·米科瓦伊奇克(Krystian Mikolajczyk)及科迪利亚·施密德(Cordelia Schmid)在“局部描述符的性能评估(A performance evaluation of local descriptors)”(关于模式分析及机器智能的电气与电子工程师协会会刊(IEEE Transactions on Pattern Analysis and Machine Intelligence),10,27,1615-1630页,2005)中的参考论文中描述这些描述符。

[0024] 步骤102中参考图像的选择及所述参考图像中显著的ROI或对象的识别可由用户提供。注意所述显著的所关注的区域或对象将通常为所述图像中的唯一对象。然而,参考图像的选择可自动完成(如图3中所展示)。参考图3,通过使用经计算的特征点,可使用步骤201将一组图像识别为类似于所述参考图像。相似性测量的实例是直方图相交,其计算两个图像之间的颜色直方图差。特定来说,可通过使用众所周知的聚类算法(例如,K平均聚类算法或ISODATA)来实现此步骤(参见下文参考)。随后在步骤202中,从所述经识别的一组类似图像中的一者选择参考图像。这可通过使用若干基于图像的标准(例如,图像质量、图像合成及图像美学)以从所述经识别的一组类似图像选择最佳图像来实现。最后在步骤203中,可由用户提供显著的ROI或对象的识别。

[0025] 使用由劳(Lowe)描述的方法来匹配SIFT特征产生参考图像与目标图像之间的一组经匹配的关键点。然而,在杂乱的场景(例如,消费者图像)中,错误匹配是相当普遍的。错误匹配在经匹配的点不对应于两个场景中的相同对象时发生。为去除错误匹配,对步骤103中的所述经匹配的点强加若干约束。特定来说,加拉格尔(Gallagher)等人在以下专利中描述用于去除错误匹配的方法:2012年4月3日颁布的第8,150,098号美国专利“通过位置对图像进行分组(Grouping images by location)”。作为第一步骤,所述经匹配的点基于所述点之间的2D欧氏距离而被集群到每一图像中的空间群组中。在优选实施例中,使用ISODATA聚类算法。此算法的快速实施是可行的(参见N.麦玛萨德奇(N.Memarsadeghi)等人在2006年的国际计算几何及应用期刊(International Journal of Computational Geometry and Applications,2006)中的论述)。消除离群点及具有小于特定数目的点(例如,5个)的集群。接着,使剩余经集群的关键点通过滤波器,所述滤波器应用旨在去除错误匹配的若干约束。由于同一对象上的关键点在空间上被约束到由所述对象占据的区域,因此我们预期参考图像中的集群将对应于针对正确匹配的目标图像中的集群。针对参考图像中的每一集群而确定关联得分,所述关联得分经计算作为在与目标图像中的集群中的点相匹配的此集群中的点的比例,此集群中的点具有最强匹配(即,最大数目的点与此集群相匹配)。消除具有低于阈值(例如,0.5)的关联得分的集群。下一个约束确保从参考到目标的点的全局轨迹是一致的,即,场景中的所有对象在大体上相同方向移动。对于参考图像中的所有剩余点,计算针对参考图像与目标图像之间的点的轨迹的平均方向。消除偏离超过此平均方向(例如,1.5a)的特定标准偏差(a)的点。接下来,计算参考图像中的每一集群的形心及消除落在此形心的特定标准偏差之外的点。此约束消除集群内的空间离群点。如果在过滤过程之后剩余关键点,那么这就指示参考图像与目标图像之间的匹配。留下的关键点的数目越大,匹配就越可靠,因此匹配的可能性与剩余匹配关键点的数目成比例。图4中描绘具有匹配关键点的两个图像的实例。在此情形中,挂在墙上的图片在两个图像之间是相同的,且因此关键点展示为良好匹配,尽管所述图像中的一者中的图片被部分遮挡。

[0026] 接下来在步骤104中,在经识别的一组类似图像中的一者中指定显著的ROI或对象的参考位置。这可由用户经由图形用户接口交互地完成。举例来说,ROI的参考位置可为如图4中所展示的墙上的图片的位置。

[0027] 在步骤105中,将经识别的图像中的每一者的对应ROI或对象与如步骤104中所指定的显著的ROI或对象的参考位置对准。对准的类型可涉及各种几何图像变换,其包含(但不限于)旋转、移位、按比例调整、缩放、裁剪、透视变换或其组合。通过确定图像内的经匹配的ROI或对象的关键点的像素坐标来实现这种情况。举例来说,可通过只覆盖参考图像中的参考ROI的所有关键点来确定边界框。注意所述边界框的实际坐标应按照图像的分辨率(即,宽度及高度)来正规化,使得可独立于图像的分辨率及高宽比来实现匹配。此参考边界框可随后用以界定参考坐标。参考图4,实例参考边界框被指示为左边参考图像上的虚线白框。针对图像集合中的所有经匹配的图像重复此过程。最后,通过将相应边界框与参考边界框对准来几何变换所有经匹配的图像。在图4的情形中,由于两个边界框的位置相当一致地对准,因此将无需此对准。另一类型的对准包含使用数字图像处理操作(例如,色彩平衡、色阶调整、亮度增强、对比度增强、噪声调整或其组合)将经检测的图像与参考图像光度对准。另一类型的对准包含将对应于特殊效应的滤波器应用到经识别的图像或参考图像。

[0028] 在对准之后,在步骤106中,将经检测的图像按时间方式(或以其它方式,例如按人物的数目或人物的年龄)进行排序。图5说明若干情景,但本发明不由此处所描绘的这些情景来限制。图5A说明按捕获时间的排序情景。为按时间顺序来排序经匹配的图像,可使用从图像的标头(例如, JPEG图像的EXIF标头)提取的捕获日期/时间信息。在图5B中,按照在图像中呈现的人物的数目来完成排序。这可通过使用人脸检测算法来实现,例如,“用于对象辨识的局部外观及空间关系的概率建模(Probabilistic Modeling of Local Appearance and Spatial Relationships for Object Recognition)”(H.施耐德曼(H.Schneiderman)及T.堪纳德(T.Kanade),计算机视觉与模式识别(CVPR)'98学报,45-51页)中所描述的人脸检测器。图5C说明按年龄的排序情景。这可通过使用人脸检测及聚类算法的组合来实现。将人脸检测算法应用于首先确定经检测的人脸的位置。随后(例如)可使用如由T.F.库茨(T.F.Cootes)、C.J.泰勒(C.J.Taylor)、D.H.库珀(D.H.Cooper)及J.格雷汉姆(J.Graham)在论文“主动形状模型-其训练及应用(Active shape models-their training and application)”(计算机视觉及图像理解(Computer Vision and Image Understanding)(61):38-59,1995)中描述的主动形状模型来提取面部特征以提取面部特征点。这些面部特征点可随后用以确定具有类似人脸(即,具有类似面部特征点的人脸)的集群。人脸的年龄随后可由用户简档数据(如果可用)来确定或由年龄及性别算法(例如,在由M.-H.杨(M.-H.Yang)及B.蒙哈达(B.Moghaddam)所著的“用于视觉性别分类的支持向量机(Support vector machines for visual gender classification)”(模式辨识国际会议学报(Proc.ICPR),2000)及由X.耿(X.Geng)、Z.-H.周(Z.-H.Zhou)、Y.张(Y.Zhang)、G.李(G.Li)及H.戴(H.Dai)所著的“从面部衰老模式学习自动年龄估计(Learning from facial aging patterns for automatic age estimation)”(2006年国际多媒体大会(ACMMULTIMEDIA 2006))中所描述的那些算法)来估计。图5D说明按时间顺序对具有同一人物的图像的排序情景。可再次通过使用人脸检测及集群以识别同一人物的图像来实现此种情况,随后从图像标头提取日期/时间信息以确定先后顺序。应注意,在本发明中可使用非人物的其它语义

内容的对象。举例来说,频繁反复出现的对象(例如,汽车、宠物、玩具、植物)可用于各种排序情景而不限于此处所描述的情景。

[0029] 最后,在步骤107中,使用经排序的一组图像来创建演示或可操纵视图。举例来说,图5中的白色箭头说明借助于上文所描述的各种情景中的类似背景图片来呈现或操纵图像的顺序。图6提供对整体应用的更详细说明。参考图6,顶部图像为参考图像,其中参考对象为墙上的图片。位于左侧的下部图像为原始图像中的一者,其经检测为具有与墙上的图片相同的图片。在对准步骤105(在此情形中通过缩放及裁剪操作)之后,位于右侧的下部图像(图像2)相对于墙上的图片更加一致地与位于顶部的参考图像(图像1)对准。当按连续顺序在幻灯片中使用时(即,图像1由图像2跟随)时,所得演示将更为赏心悦目,这是由于当从一个图像过渡到另一图像时类似背景看起来是静止的或较不易使观众分散注意力。此外,可添加各种过渡效果(例如,渐隐及擦除)以进一步增强整体演示。

[0030] 多媒体演示及操纵的另一实施例利用用于全景拼接及虚拟环境创建的经识别的反复出现的静止对象以及用于在经创建的虚拟环境中的布置的经提取的非静止前景对象。可按时间及空间标准(例如,先后顺序、集群、及相对物理位置、对象或人物的年龄)来对所述经提取的前景对象进行排序及组织。

[0031] 以引用的方式并入由民宇·帕克(Minwoo Park)、亚历山大·路易(Alexander Loui)及玛丽云基·库马尔(Mrityunjay Kumar)所著的标题为“识别数字图像中的高显著性区域(Identifying high saliency regions in digital images)”的美国专利8401292来描述自动场景分割及对象提取技术。简单地说,‘292专利教导一种用于通过自动分析数字图像以将其分割成多个经分割的区域来识别所述数字图像中的高显著性区域的方法,其中每一经分割的区域包含一组图像像素且通过分割边界来定边界。每一经分割的区域的显著性值通过以下步骤来确定:

[0032] a. 确定包含环绕所述经分割的区域的一组图像像素的环绕区域,其中所述环绕区域具有由所述分割边界所界定的外边界及内边界;

[0033] b. 分析所述经分割的区域中的图像像素以确定一或多个经分割的区域属性;

[0034] c. 分析所述环绕区域中的图像像素以确定一或多个对应环绕区域属性;以及

[0035] d. 响应于所述一或多个经分割的区域属性与所述对应环绕区域属性之间的差异来确定区域显著性值。

[0036] 随后响应于确定满足一或多个规定的合并标准来合并共享共同边界的相邻经分割的区域,其中规定的合并标准包含以下标准:经合并的经分割的区域的显著性高于相邻经分割的区域的显著性。最后,具有最高区域显著性值的经分割的区域中的一或多个者被指定为高显著性区域。此方法具有以下优点:由若干统计上不均匀区域组成的显著的对象可被检测为整体。其具有以下额外优点:其可有效地检测大显著的区域及小显著的区域两者。

[0037] 此外,可将各种经创建的虚拟环境呈现为一系列同心带或球,其中位于中心的“观察者”具有对场景及对象的独立控制。举例来说,用户可使用由摄像机监视的手势、指向装置或触摸屏来将带向左或向右平移,通过向前及向后行进而移动通过带。所选择的“背景带”可为连续的或一系列的重建场景,其具有填充有场景相关的经提取对象的“前景带”。可使用图2的步骤101到103来识别用于重建的背景场景。可由任何现有技术(例如,经由图切割的能量最小化)来实现场景重建。在由V.柯尔莫哥洛夫(V.Kolmogorow)及杂宾(Zabih)所

著的标题为“经由图切割的多摄像机场景重建 (Multi-camera scene reconstruction via graph cuts)”的论文(2002年5月的欧洲计算机视觉会议 (European Conference on Computer Vision, May 2002))中描述一种此方法。

[0038] 在典型场景内存在图像对象类型的四个主要类别,其包含前景、背景、静止及可移动。通过提供具有静止或固定对象(如经装饰的房间、墙壁、地板、艺术墙、花园、自然景观、建筑、城市景观及类似物)的更连续的室内及室外背景环境,这将允许用户以更自然的方式浏览图像集合。提取相关移动前景对象(如人物、车辆、玩具、动物)且允许其独立于相关背景而移动。使用反复出现的静止对象(如关键点)提供使用多个背景场景来“填充”由经提取的对象引起的任何图像数据“洞”的能力。

[0039] 移动前景对象还可经修改以包含简单动画,例如,借助于通过提供移动且在膝及肘关节处弯曲的四肢的经提取的人物的图像,或借助于与汽车朝向的方向一致的向左或向右横向移动的机动车的图像,车轮也可经独立地提取且旋转以进一步增强车辆的运动错觉。当汽车的前部或背部图像为可见时,可通过放大或缩小使汽车看起来似乎正在移动,从而给予运动错觉。可由用户使用标准图像动画工具(例如,GIMP (GNU图像处理程序 (GNU Image Manipulation Program))或Adobe边缘动画CC (Adobe Edge Animate CC))来手动创建这些简单动画。替代地,借助于经语义识别及经提取的对象,可自动创建并向用户呈现动画选项,例如,挥手、向左行走、向右行走、跳舞,等等。此外,在用户操纵虚拟环境时,经提取的对象在其进入用户的视野时可自动变成动画。

[0040] 当用户缩放及平移通过同心带或球时,前景及背景内容可独立地移动,使得用户可独立地选择背景环境及前景对象。这将允许用户创建及体验虚拟环境或使用此相同技术来容易地创建适用于打印及共享的填充场景。

[0041] 尽管已特定参考本发明的某些优选实施例来详细描述本发明,但应理解,在本发明的精神及范围内可实现变型及修改。

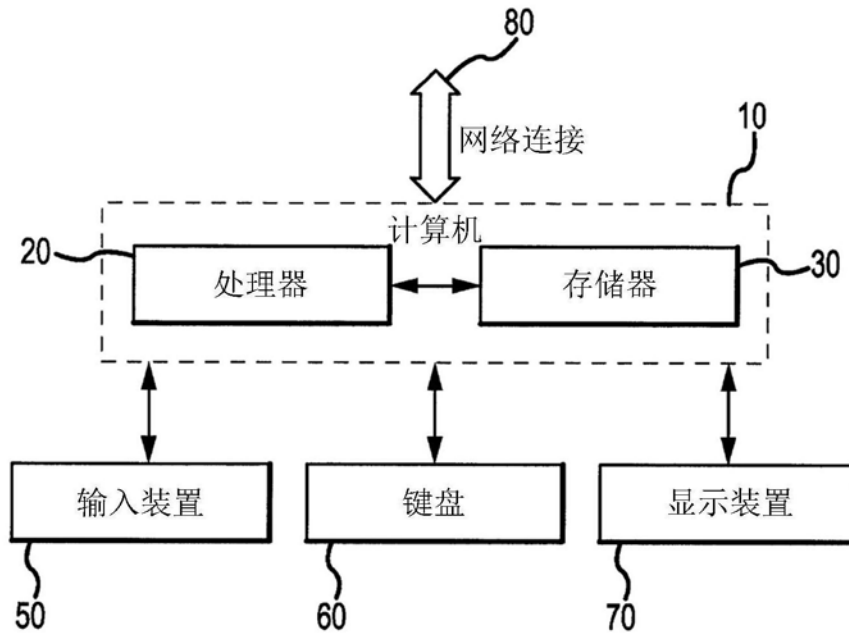


图1

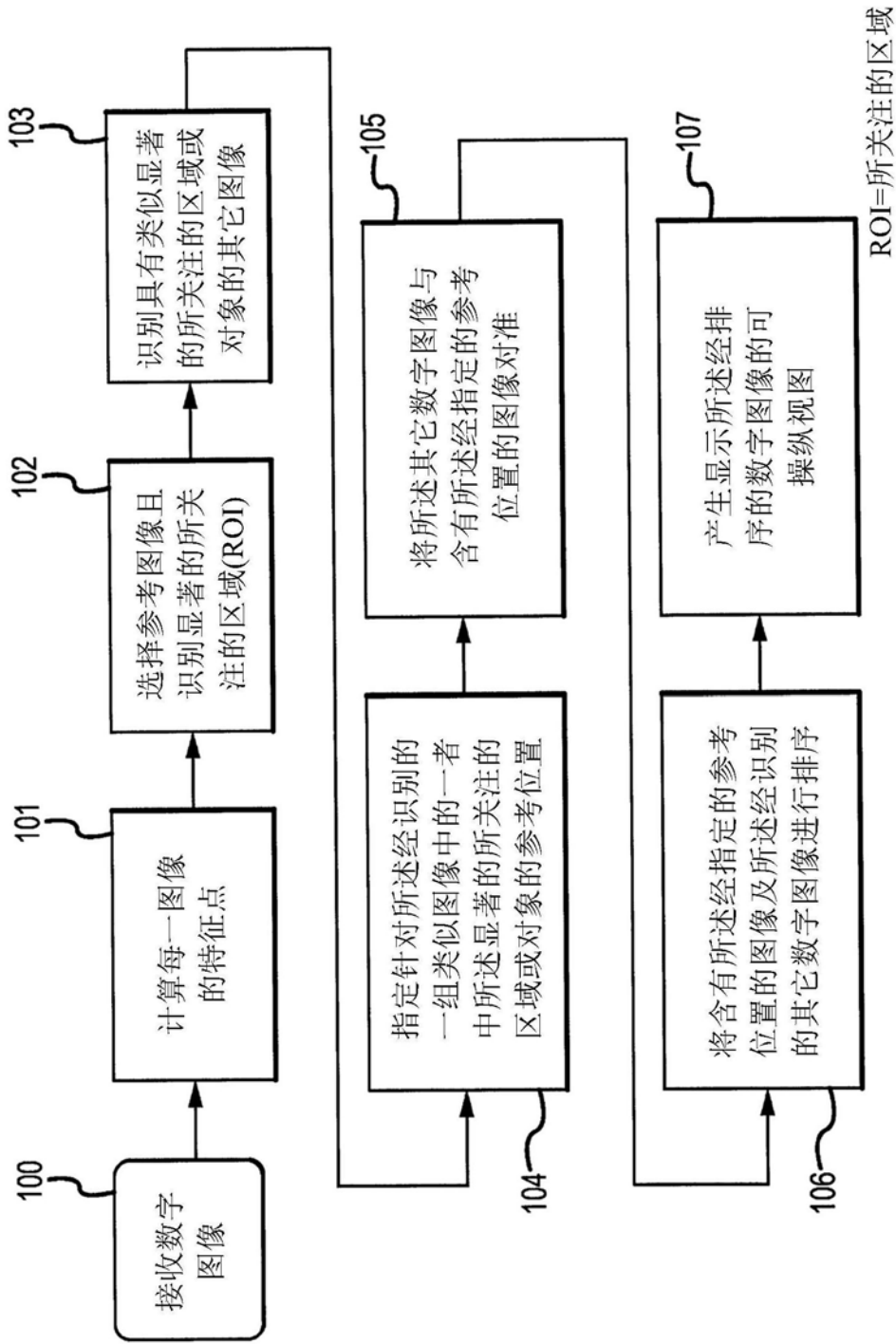


图2

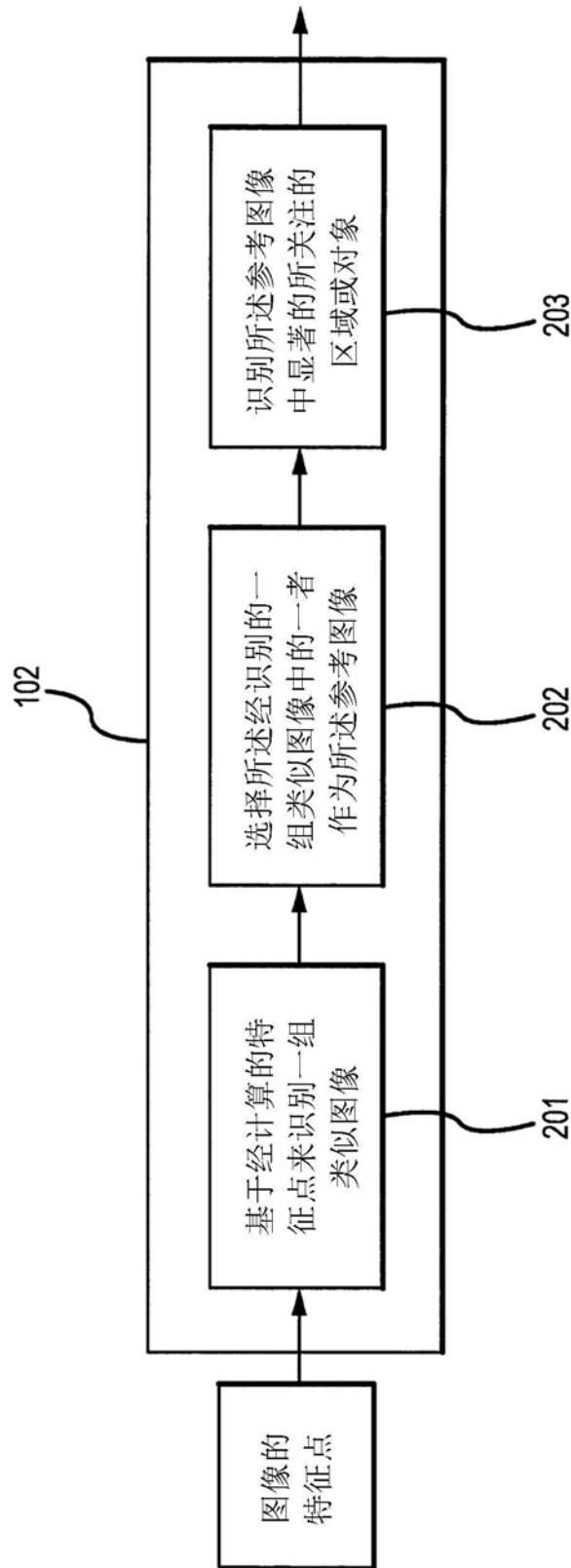


图3

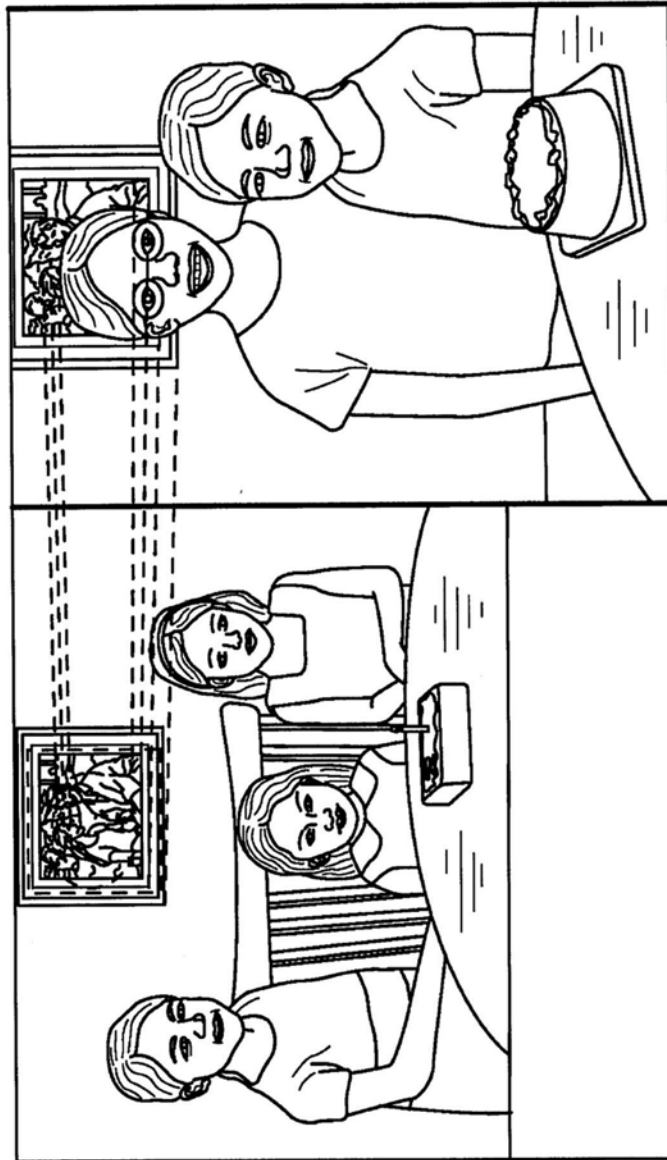


图4

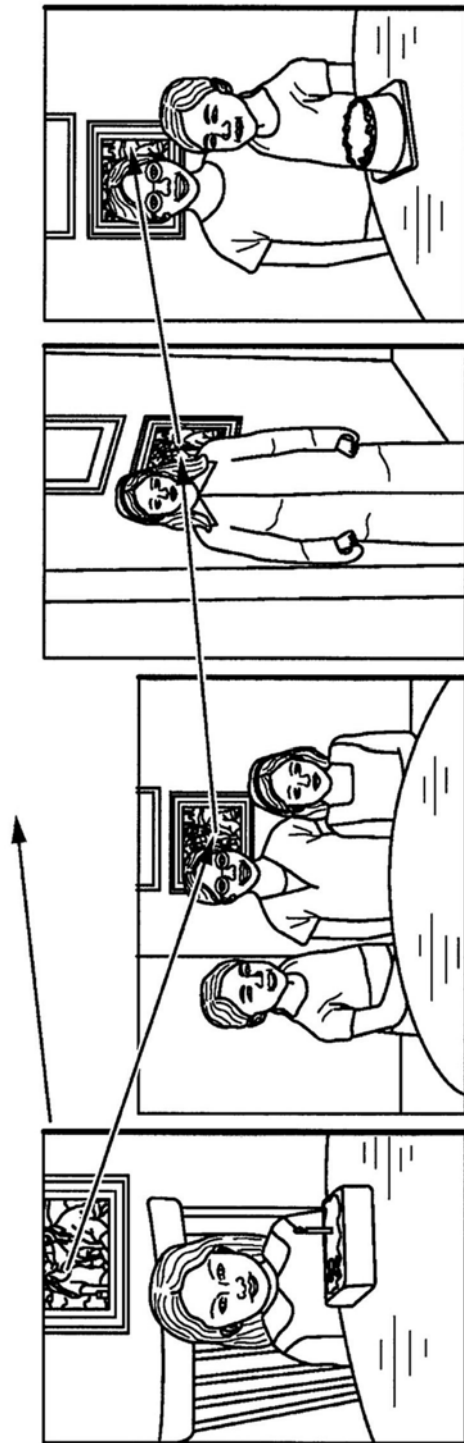


图5A



图5B

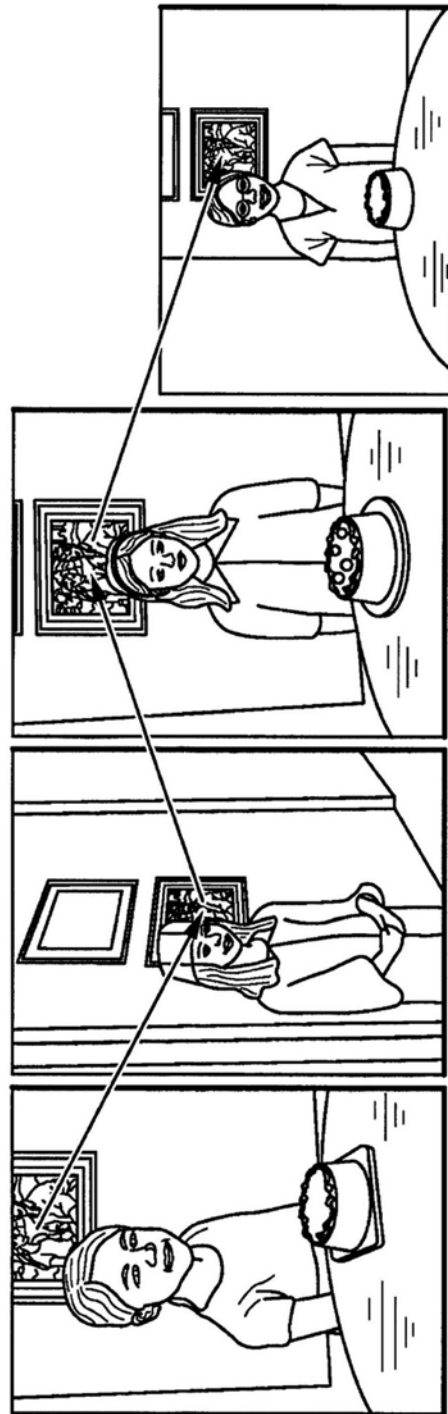


图5C

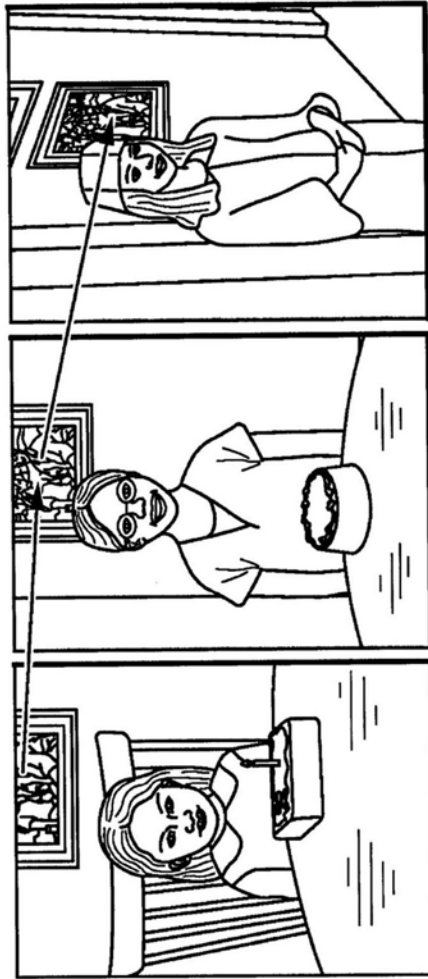


图5D

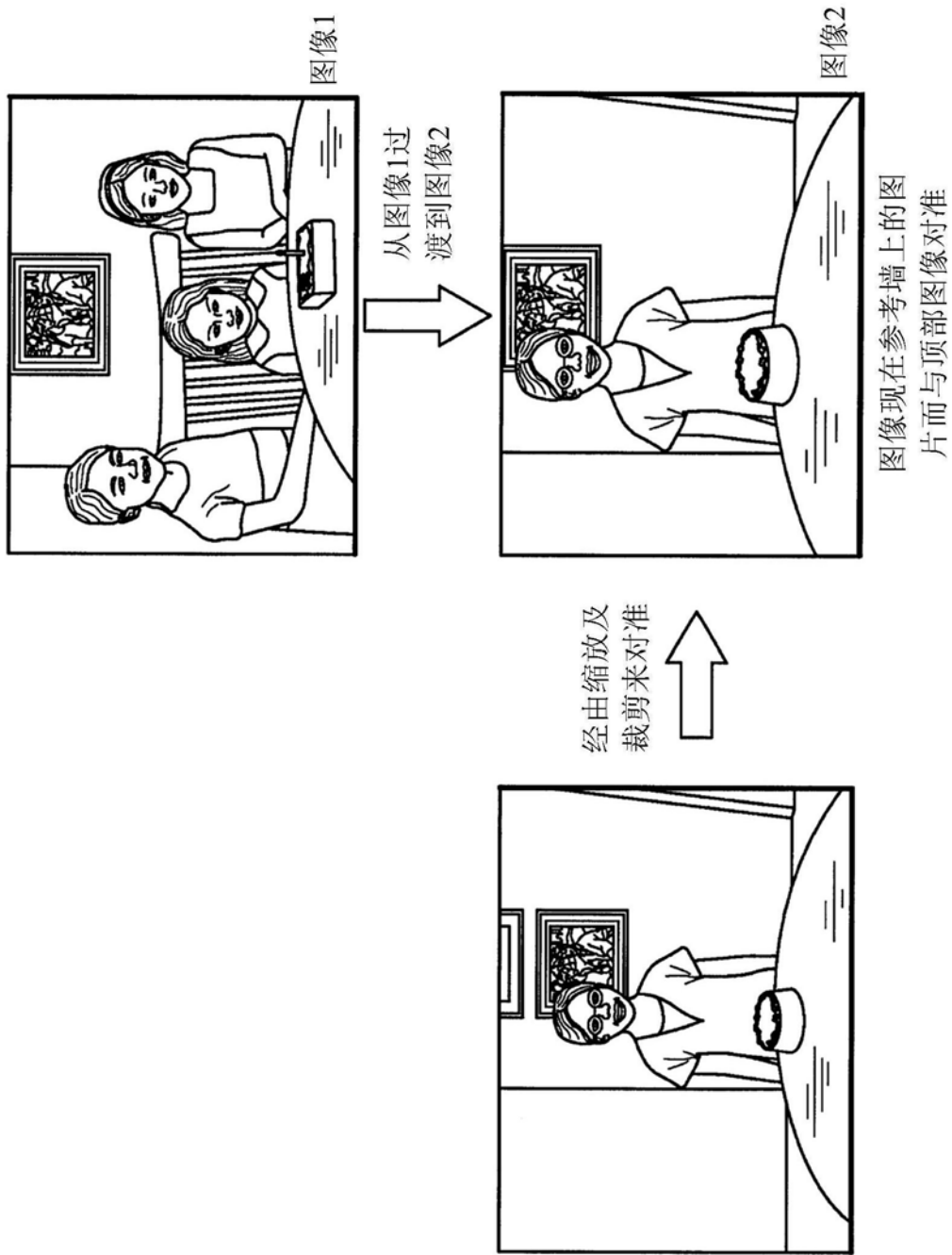


图6