

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04N 13/00 (2006.01)

H04N 7/18 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200910077230.7

[43] 公开日 2009年7月22日

[11] 公开号 CN 101489150A

[22] 申请日 2009.1.20

[21] 申请号 200910077230.7

[71] 申请人 北京航空航天大学

地址 100083 北京市海淀区学院路 37 号

[72] 发明人 赵沁平 周 忠 刘寒冰 常雪枫
吴 威

[74] 专利代理机构 北京科迪生专利代理有限责任
公司

代理人 成金玉 卢 纪

权利要求书 2 页 说明书 6 页 附图 2 页

[54] 发明名称

一种虚实混合的远程协同工作方法

[57] 摘要

一种虚实混合的远程协同工作方法，由三个拼接显示的显示器组成显示装置，每个显示器前面以 45 度角倾斜放置半透半反镜，镜子上方摄像头的镜头竖直向下，通过采集镜子反射的光线采集到用户的正对视线。在进行采集时，初始化阶段三个摄像机采集背景图像，然后采集背景不变用户进入的图像，对两组图像分别求差值，得到用户的前景图像；把三路图像进行编码，合成为一个用户半侧面的连续图像，该图像是半三维的有立体感；对图像进行压缩编码并传输，远端接收到压缩合成的前景图像之后设 alpha 通道，通过 alpha 混合把前景影像加入到三维场景中，用户可以控制自己的影像在三维场景中漫游，也可以通过特定手势与三维对象进行交互。本发明具有真实、便利、共享、方便交流等优点。



1、一种虚实混合的远程协同工作方法，其特征在于步骤如下：

第一步，构建基于多摄像机采集的远程沉浸协同工作装置，所述的工作装置包括：三个摄像机、三个呈 α 角度放置的显示器和分别位于每个显示器前以 45 度角放置的用于改变光路的半透半反镜，三个摄像机分别固定在半透半反镜正上方，镜头在半透半反镜中的投影与用户正对显示器时眼睛在半透半反镜中的投影重合，通过采集半透半反镜反射的光线采集到用户的前景影像，其中 α 大于 90 度；

第二步，利用第一步构建的装置实现三维场景中用户的实时运动影像的采集、提取、传输、加载，具体如下：

(1) 利用三个摄像机采集用户的三路视频，从视频中提取出用户的前景影像；

(2) 把三路采集的前景影像进行编码合成，生成一个用户半侧面的连续图像，该图像是半三维前景影像，有立体感；

(3) 把步骤 (2) 生成的半三维前景影像经过压缩，传输到远程节点端；

(4) 远程节点端接收到压缩的半三维前景影像之后，为其标志 alpha 通道，通过 alpha 混合把半三维前景影像加载到三维场景中，得到真人进入三维场景的效果；

(5) 控制用户影像在三维场景中漫游及交互。

2、根据权利要求 1 所述的虚实混合的远程协同工作方法，其特征在于：所述的步骤

(1) 中的提取出用户的前景影像的步骤如下：

a: 摄像机标定：通过对所述的三个摄像机进行标定；三个摄像机接收外部主机发出的同一个外触发信号，通过采集所述的半透半反镜反射的光线对同一个场景进行同步真实图像的采集；

b: 进行背景图像采集：所述的三个摄像机在同一个硬件触发信号控制下同步采集目标物体的图像，分别得到图像 ImL 、 ImM 和 ImR ；

d: 进行以用户为前景的图像采集：用户进入所述的三个摄像机的采集范围，所述的三个摄像机同步采集以原背景为背景，以用户为前景的三个图像 $Im1$ 、 $Im2$ 和 $Im3$ ；

e: 计算前景图像：对步骤 (2) 和步骤 (3) 得到的两组图像 ImL 和 $Im1$ 、 ImM 和 $Im2$ 及 ImR 和 $Im3$ 分别求差值，得到用户的前景图像。

3、根据权利要求 1 所述的一种虚实混合的远程协同工作方法，其特征在于：所述的三个摄像机的光学参数完全相同，且三个摄像机镜头均竖直向下，与半透半反镜呈 45

度夹角，主轴平行。

4、根据权利要求1所述的虚实混合的远程协同工作方法，其特征在于：所述的通过交互装置控制用户影像在三维场景中漫游及交互的，控制装置有键盘鼠标方式、方向盘脚踏方式和手势方式三种，可择其一进行，或同时选择。

5、根据权利要求4所述的虚实混合的远程协同工作方法，其特征在于：所述的约定手势至少有四种：单手向右挥、单手向左挥、双手向前推和双手向回拉，分别依次代表的含义依次是：视角向左转，视角向右转，影像向后退，影像向前进。

6、根据权利要求4所述的一种虚实混合的远程协同工作方法，其特征在于：所述的约定手势还包括两种，双手张开表示对当前视野进行放大，双手合拢表示对当前视野进行缩小。

一种虚实混合的远程协同工作方法

技术领域

本发明涉及一种虚实混合的远程协同工作方法，用于构建真人共享三维场景进行协同工作的虚拟工作区，属于虚拟现实领域。

背景技术

目前用户参与的共享交流是虚拟现实领域研究的热点之一，已广泛应用于游戏、共享工作空间、视频会议等各个领域。采用的方法主要有三种。

第一种，marker方法。该方法通过对人体关节点增加标志点达到对人体跟踪的目的，需要特定的设备。marker设备从实现方式上又可以分为两大类：磁性跟踪设备和光学跟踪设备。前者用一个接收器接收人身体佩戴的marker发射出的电磁频率，根据不同marker频率的不同标定关节点的位置。后者通过在人体关节点安装光发射装置或反射装置，通过被动接收装置或主动式光学捕获系统实现对关节点的标识。然后把关节点的运动映射到模型上，用真人驱动模型在虚拟环境中与其他用户交互。

第二种，模型方法。该方法广泛应用于动作类游戏中。用一个角色模型代表玩家，玩家通过键盘鼠标控制该模型完成在三维场景中与其他用户的交互体验。

第三种，视频方法。该方法主要被应用于视频会议中，采集用户的视频，在各个参与节点之间进行点对点传输，每个参与用户看到的也是其他用户的视频。

综上所述，在本发明出现之前，用户参与的共享交流要么采用模型的方法，要么采用互传视频的方法，前者虽然能够很好地与虚拟环境进行交互，但是模型不够真实。而视频的方法无法与虚拟环境结合，无法达到“共享”的目的。虽然近年来也出现了一些真人进入的远程沉浸系统，但它们有的是通过多角度视频生成有深度的像素点集合进行传输再重新组合，有的是通过前景提取的方式抠出人体影像再叠加到虚拟场景中，但前者由于数据量大，只能采用不可靠的方式传输，造成丢包，所以容易造成人体影像的空洞，而后者仅是二维的视频影像与三维场景的简单叠加，既不能进行有效交互也没有可供讨论的道具，因此其效果只是类似于视频会议，并无法体现虚实融合的优势。

发明内容

本发明的技术问题：克服现有方法的不足，提供一种虚实混合的远程协同工作方法，该方法具有真实、便利、共享、方便交流等优点；且还可以控制自己的影像在三

维场景中漫游，也可以通过特定手势与三维对象进行交互。

本发明的技术方案：一种虚实混合的远程协同工作方法步骤如下：

第一步，构基于多摄像机采集的远程沉浸协同工作装置，所述的工作装置包括：三个摄像机、三个呈 α 角度放置的显示器和分别位于每个显示器前以 45 度角放置的用于改变光路的半透半反镜，三个摄像机分别固定在半透半反镜正上方，镜头在半透半反镜中的投影与用户正对显示器时眼睛在半透半反镜中的投影重合，通过采集半透半反镜反射的光线采集到用户的前景影像，其中 α 大于 90 度；

第二步，利用第一步构建的装置实现三维场景中用户的实时运动影像的采集、提取、传输、加载，具体如下：

(1) 利用三个摄像机采集用户的三路视频，从视频中提取出用户的前景影像；

(2) 把三路采集的前景影像进行编码合成，生成一个用户半侧面的连续图像，该图像是半三维前景影像，有立体感；

(3) 把步骤 (2) 生成的半三维前景影像经过压缩，传输到远程节点端；

(4) 远程节点端接收到压缩的半三维前景影像之后，为其标志 alpha 通道，通过 alpha 混合把半三维前景影像加载到三维场景中，得到真人进入三维场景的效果；

(5) 控制用户影像在三维场景中的漫游及交互。

所述的步骤 (1) 中的提取出用户的前景影像的步骤如下：

a: 摄像机标定：通过对所述的三个摄像机进行标定；三个摄像机接收外部主机发出的同一个外触发信号，通过采集所述的半透半反镜反射的光线对同一个场景进行同步真实图像的采集；

b: 进行背景图像采集：所述的三个摄像机在同一个硬件触发信号控制下同步采集目标物体的图像，分别得到图像 I_{mL} 、 I_{mM} 和 I_{mR} ；

d: 进行以用户为前景的图像采集：用户进入所述的三个摄像机的采集范围，所述的三个摄像机同步采集以原背景为背景，以用户为前景的三个图像 I_{m1} 、 I_{m2} 和 I_{m3} ；

e: 计算前景图像：对步骤 (2) 和步骤 (3) 得到的两组图像 I_{mL} 和 I_{m1} 、 I_{mM} 和 I_{m2} 及 I_{mR} 和 I_{m3} 分别求差值，得到用户的前景图像。

所述的三个摄像机的光学参数完全相同，且三个摄像机镜头均竖直向下，与半透半反镜呈 45 度夹角，主轴平行。

所述的通过交互装置控制用户影像在三维场景中漫游及交互的，控制装置有键盘鼠标方式、方向盘脚踏方式和手势方式三种，可择其一进行，或同时选择。

所述的约定手势至少有四种：单手向右挥、单手向左挥、双手向前推和双手向回拉，

分别依次代表的含义依次是：视角向左转，视角向右转，影像向后退，影像向前进。

所述的约定手势还包括两种双手张开表示对当前视野进行放大，双手合拢表示对当前视野进行缩小。

本发明与现有技术相比的优点在于：

(1) 本发明主要利用视频比模型真实感好的特性，通过对用户的视频进行前景提取和与虚拟场景的alpha混合使真人影像进入到虚拟场景中；为了克服二维视频放在三维场景中不真实的缺点，对三路视频进行压缩合成，生成了有侧面连续信息的半三维影像；为了克服视频无法与虚拟对象进行交互的缺点，采用对特定手势的识别来完成交互，这种使真人进入的虚实混合远程协同设计方法具有真实、便利、共享、方便交流等优点。

(2) 使真人影像进入到三维场景，通过网络传输和远程节点在共享的三维场景中进行面对面的交互，在三维场景中，加入的是人体的半三维影像，在视角发生一定范围偏转的时候，可以看到用户侧面的连续影像，具有一定的三维效果，真实感强；在三维场景中，用户可以控制自己的位置和视角，实现在虚拟场景中的自由活动，而且可以通过对特定手势的识别实现对虚拟场景的正向交互。

附图说明

图 1 为本发明的虚拟协同工作台装置图；

图 2 为本发明的真人进入三维场景方法的流程图；

图 3 为本发明的人体影像前景提取方法流程图；

图 4 为本发明的半三维影像生成方法流程图。

具体实施方式

本发明通过如图 1 所示的基于多摄像机采集的远程沉浸协同工作装置进行采集和显示，在虚拟场景中进行协同工作的人是真人的影像，其全部表情、动作都是实时采集的真人视频。

图 1 中的装置由三块半透半反镜 2、三个摄像机 3 和三个显示器 1 组合。三块半透半反镜 2 通过一个固定架 4 固定或通过三个固定架分别固定，固定架 4 竖直的一面与显示器相邻，倾斜的一面放置半透半反镜 2，每个半透半反镜 2 与显示器 1 的夹角为 45 度，相邻两个显示器 2 间的夹角为 α ，其中 α 大于 90 度，因此半透半反镜不能切割为矩形，假设显示器 1 的尺寸为 $a*b$ （长 \times 高），则半透半反镜 2 所切割成的梯形为一个等腰梯形，它的上底（较长的一端）为 a ，下底（较短的一端）为 $a-2b*\cos(\alpha/2)$ ，两腰长度为 $1.414b$ 。

三个摄像机的光学参数完全相同，且三个摄像机分别固定在半透半反镜 2 的正上

方，镜头竖直向下，与半透半反镜呈 45 度夹角，主轴平行，镜头在半透半反镜中的投影与用户正对显示器时眼睛在半透半反镜中的投影重合。在摄像机镜头所在平面放置深色（黑色或深蓝色等）挡光板，同时挡光板上开一个小孔，孔的尺寸和摄像机镜头尺寸相当，摄像机通过该小孔采集图像。

如图 2 所示，本发明一种虚实混合的远程协同工作方法的具体步骤为：

(1) 采集用户的三路视频，从视频中提取出用户的前景影像；

其中人体前景影像提取的具体实施方法如图 3 所示，具体步骤如下：

a: 摄像机标定：先把三个摄像机镜头向下固定在架子上，让三个摄像机共同采集一条竖直向下的线（比如两面墙相交的缝），校正摄像机，使摄像机采集的线的图像和铅锤线重合，从而保证方向竖直并且主轴之间互相平行；

b: 进行背景图像采集：三个摄像机在同一个硬件触发信号控制下同步采集目标物体的图像，分别得到图像 ImL 、 ImM 和 ImR ，这里的 ImL 、 ImM 和 ImR 分别是一个像素结构的数组，该像素结构是一个三元组 $\{R, G, B\}$ ，元组中 R, G, B 分量分别代表红，绿，蓝的颜色值，范围是 $0 \sim 255$ ，而 ImL 、 ImM 和 ImR 就是这个三元组的一个数组，数组的大小就是采集图像的像素数。这里的 ImL 、 ImM 和 ImR 不是实际采集的图像，而是通过计算平均值得到的。由于光线扰动等因素的干扰，采集背景图像的时候，通常采用一个数组，存储一段时间采集的背景不变的图像（如 10 帧），然后对每个像素的值取平均，从而得到均一化的 ImL 、 ImM 和 ImR 。以 ImL 某一个像素 $ImL_{i,j}$ （ i 和 j 分别表示该像素在整个图像中的行和列）的 R 分量为例，就是用摄像机采集 10 帧，分别记为 $ImL[1]$ 、 $ImL[2]$ 、 \dots 、 $ImL[10]$ ，则像素 $ImL_{i,j}$ 的 R 分量分别记为 $ImL[1]_{i,j}.R$ 、 $ImL[2]_{i,j}.R$ 、 \dots 、 $ImL[10]_{i,j}.R$ ，则 $ImL_{i,j}.R = (ImL[1]_{i,j}.R + ImL[2]_{i,j}.R + \dots + ImL[10]_{i,j}.R) / 10$ 。按照这个方法对其他分量和每一个像素进行计算就得到了整个背景图像的均一化值。

c: 进行以用户为前景的图像采集：用户进入三个摄像机的采集范围，三个摄像机同步采集以原背景为背景，以用户为前景的三个图像 $Im1$ 、 $Im2$ 和 $Im3$ ；

d: 计算前景图像：对步骤 (2) 和步骤 (3) 得到的两组图像 ImL 和 $Im1$ 、 ImM 和 $Im2$ 及 ImR 和 $Im3$ 对应的像素的对应分量分别求差值，超过阈值的即为用户的前景图像。

(2) 把三个角度采集的前景图像进行编码合成，生成一个用户半侧面的连续图像，该图像是半三维的，有一定的立体感；

图像合成的具体实施方法如图 4 所示，具体步骤如下：

a: 投影变换，由于用户的前景图像是在不同角度下拍摄的，它们并不在一个投影平面上，如果直接对重叠图像进行无缝拼接就会破坏实际景物的视觉一致性，因此，结

合本方法的应用范围，需要将得到的反映各自投影平面的图像映射到一个标准投影——柱面投影上。

b: 特征点提取，在选择特征点的时候，首先要保证旋转、缩放的不变性；其次特征点的数量应该丰富，因为只有这样，才能充分表达出图像之间的关联关系；另外就是特征点的提取速度要快，因为本方法是要应用于实时的交互系统中，所以必须满足每秒15帧的图像性能指标。SURF特征点是可供选择的一种。

c: 特征点匹配，以SURF特征点为例，包含了每个特征点属性的矢量，所以可以采用距离测度来进行匹配，如选用欧式距离来描述两个特征点之间的相似性。 n 维空间的欧式距离公式如下：

$$d = \sqrt{\sum (x_{i1} - x_{i2})^2}$$

其中 $i=1, 2, \dots, n$ 。 x_{i1} 表示第一个点的第 i 维坐标， x_{i2} 表示第二个点的第 i 维坐标。 n 维欧氏空间是一个点集，它的每个点可以表示为 (x_1, x_2, \dots, x_n) ，其中 $x_i (i=1, 2, \dots, n)$ 是实数，称为 x 的第 i 个坐标，同样也有 $y = (y_1, y_2, \dots, y_n)$ 。两个点 x 和 y 之间的距离 $d(x, y)$ 定义为上述公式。

每个SURF特征点都有一个64位的描述矢量。将每个点的描述矢量作为64维空间的一个点，利用欧氏距离对两点进行距离测度，距离度量值越小，两点越相似。

d: 图像融合，通过上面的算法得到图像 ImX 和 ImY 的多组相匹配的特征点 x 和 y 之后，让 y 所在图像 ImY 作平移，使得 y 和 x 重合，从而实现图像的拼接，生成可以转动一定角度的半三维影像。

(3) 把生成的半三维影像进行压缩，传输到远程节点端。对半三维影像的压缩分为两个部分，一部分是对单帧的压缩，一部分是对视频流的压缩。

首先是对单帧的压缩，由于生成的半三维影像中的有效数据只有用户的前景影像，为了压缩数据，这里把原来的三元组的像素结构改成了四元组结构 $\{n, R, G, B\}$ ，其中 n 代表某一前景像素点在原图像中的位置， R, G, B 分别等于该前景像素点的 R, G, B 值。虽然元组多出一维，但是省略了大片背景的传输，还是起到了压缩的效果。另外在传输的时候还可以采用一些压缩算法对图像数据进行压缩，使数据量更小。

其次是对视频流的压缩，由于该视频流对象的位置、姿势基本不变，因此各帧间相关性很大，可以对前后两帧做插值，只传递帧间改变的部分，当改变超过一定的阈值才传递整个图像，这样就能更加节省对网络带宽的占用。

(4) 远程节点端接收到压缩合成的前景图像之后，对其解压缩，并为其标志 α

通道，即把三元组改变为四元组 $\{R, G, B, A\}$ ，图像中前景部分像素的 A 值为 1，背景部分像素的 A 值为 0。然后通过 alpha 混合把影像加入到三维场景中，实现背景透明的效果。

(5) 通过交互装置控制自己的影像在三维场景中漫游，通过对约定手势的识别与三维对象进行交互。

控制漫游的方式一共有三种，分别是键盘鼠标方式、方向盘脚踏方式和手势方式，可择其一进行，或同时选择。其中键盘鼠标方式采用一般游戏中常用的操作方法，通过鼠标控制视角的方向，通过键盘“上”、“下”、“左”、“右”（或“W”、“S”、“A”、“D”）控制本地第一人称视点位置的移动，在远程节点端看来就是该用户影像位置的移动。第二种是方向盘脚踏的方式，类似于机动车的驾驶方式，方向盘控制视角，脚踏控制本地第一人称视点位置的移动。另外一种就是手势方式，通过对约定手势的识别实现在场景中的漫游，如单手向右挥、单手向左挥、双手向前推和双手向回拉，分别依次代表的含义依次是：视角向左转，视角向右转，影像向后退，影像向前进。

还可以通过对约定手势的识别与三维对象进行交互，如双手张开表示对当前视野进行放大，双手合拢表示对当前视野进行缩小等。

本发明未详细阐述的部分属于本领域技术人员公知技术。

最后所应说明的是，以上所述仅是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明利用把真人前景影像加入三维场景中，并通过对特定手势的识别与三维场景进行交互原理的前提下，还可以做出若干改进或等同替换，这些改进和等同替换也应视为本发明的保护范围。

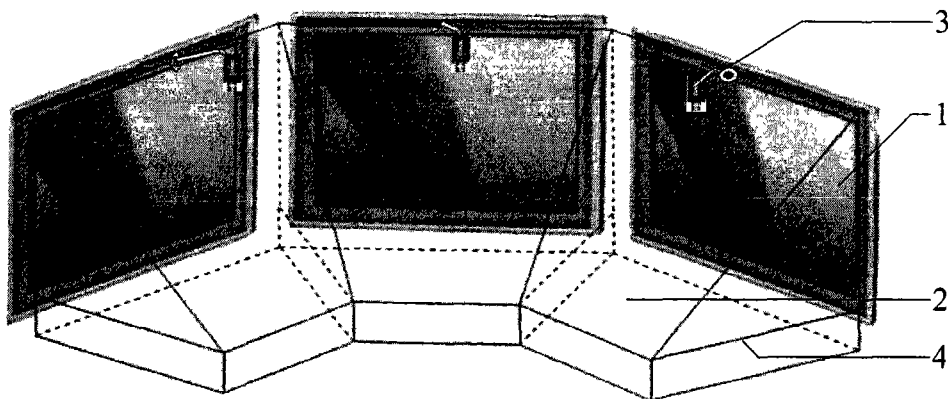


图 1

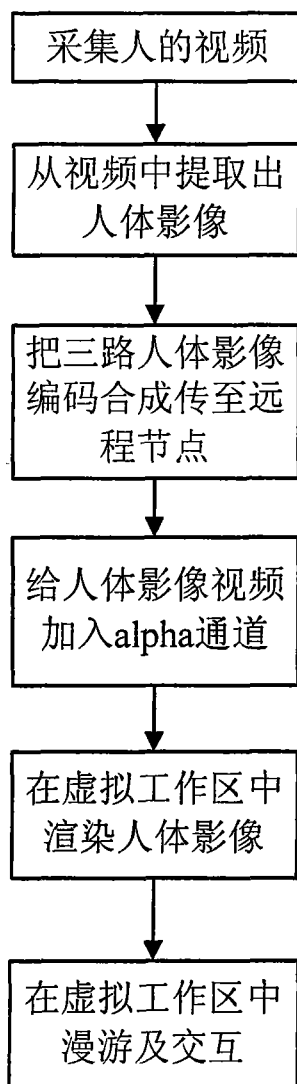


图 2

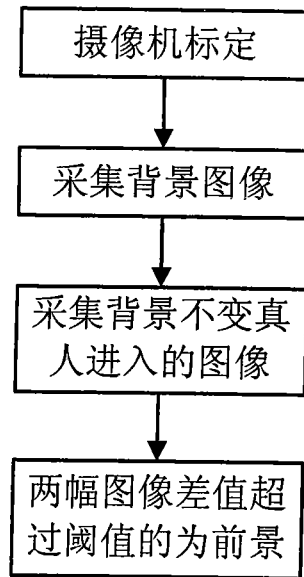


图 3

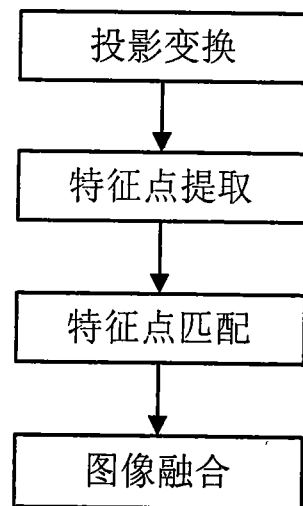


图 4