



(12) 发明专利申请

(10) 申请公布号 CN 102638653 A

(43) 申请公布日 2012. 08. 15

(21) 申请号 201210052106. 7

(22) 申请日 2012. 03. 01

(71) 申请人 北京航空航天大学
地址 100191 北京市海淀区学院路 37 号

(72) 发明人 周忠 吴威 李涛 赵沁平

(74) 专利代理机构 北京科迪生专利代理有限责
任公司 11251

代理人 许玉明 贾玉忠

(51) Int. Cl.

H04N 5/232(2006. 01)

G06T 7/00(2006. 01)

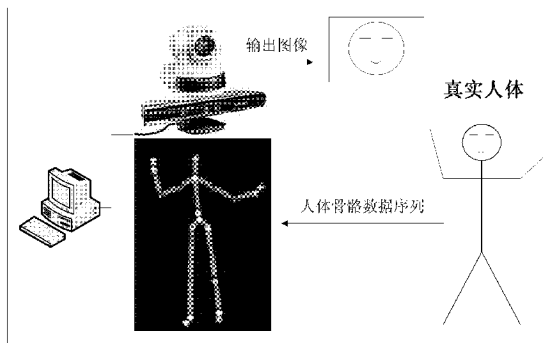
权利要求书 2 页 说明书 4 页 附图 4 页

(54) 发明名称

一种基于 Kinect 的自动面部追踪方法

(57) 摘要

一种基于 Kinect 的自动面部追踪方法,步骤为:(1) 将 Kinect 与可编程旋转摄像机组成为追踪设备并进行标定;(2) 通过 Kinect 获取人体骨骼数据序列,并从中提取头部骨骼点在 Kinect 坐标系中的坐标;(3) 利用当前摄像机的外参数和当前 Kinect 的外参数得到 Kinect 坐标系到摄像机坐标系的变换矩阵;(4) 利用步骤(3)中获取的变换矩阵将步骤(2)中得到的头部骨骼点坐标变换到摄像机坐标系中,而后计算摄像机在水平和垂直方向需要旋转的角度以及需要缩放的倍数;(5) 根据步骤(4)中求得的角度和倍数,驱动摄像机进行相应的旋转和缩放,使得摄像机能够追踪人的面部运动。本发明利用连续获取的骨骼数据序列驱动摄像机旋转和缩放,可有效地实现自动面部追踪。



1. 一种基于 Kinect 的自动面部追踪方法,其特征在于包含如下步骤:

(1) 将追踪设备放置在稳定的位置,追踪设备由 Kinect 和可编程旋转摄像机组成,通过标定得到摄像机的内参数、Kinect 的内参数以及摄像机的外参数、Kinect 的外参数;

(2) 通过 Kinect 获取人体骨骼数据序列,并从中提取头部骨骼点在 Kinect 坐标系中的坐标,提取方法为直接提取,或者是在提取过程中结合骨骼点之间的约束关系进行优化;

(3) 完成步骤 (2) 以后,利用步骤 (1) 中得到的摄像机的外参数和 Kinect 的外参数,按照在步骤 (1) 中追踪设备位置放置方式,计算当前摄像机的外参数和当前 Kinect 的外参数并利用它们得到 Kinect 坐标系到摄像机坐标系的变换矩阵;

(4) 利用步骤 (3) 中获取的变换矩阵将步骤 (2) 中得到的头部骨骼点在 Kinect 坐标系中的坐标变换到摄像机坐标系中,得到头部骨骼点在摄像机坐标系中的坐标并利用该坐标计算摄像机在水平和垂直方向需要旋转的角度以及需要缩放的倍数;

(5) 根据步骤 (4) 中求得的角度和倍数,驱动摄像机进行相应的旋转和缩放,使得摄像机能够追踪人的面部运动。

2. 根据权利要求 1 所述的基于 Kinect 的自动面部追踪方法,其特征在于:通过将 Kinect 与可编程旋转摄像机标定到同一个坐标空间中,对 Kinect 获取的人体骨骼数据序列进行处理和提取,来驱动摄像机对人的面部进行自动追踪。

3. 根据权利要求 1 所述的基于 Kinect 的自动面部追踪方法,其特征在于:所述步骤 (1) 的追踪设备中的 Kinect 和可编程旋转摄像机的位置放置方式有两种,第一种是在 Kinect 上集成或固定摄像机,第二种是将 Kinect 和摄像机分别独立放置在不同的位置。

4. 根据权利要求 1 所述的基于 Kinect 的自动面部追踪方法,其特征在于:所述步骤 (2) 中若利用骨骼点之间的约束关系进行优化提取头部骨骼点在 Kinect 坐标系中的坐标,其方法为:头部骨骼点附近的骨骼点包括肩部中心骨骼点、左肩骨骼点、右肩骨骼点和脊柱中心骨骼点,它们之间存在着一定的约束关系,肩部中间骨骼点、左肩骨骼点和右肩骨骼点认为是共面,肩部中心骨骼点、脊柱中心骨骼点认为是共线,头部骨骼点受到这二者的约束,设头部骨骼点、肩部中心骨骼点、左肩骨骼点、右肩骨骼点和脊柱中心骨骼点在 Kinect 坐标系中的齐次坐标分别为 $(H_x, H_y, H_z, 1)$ 、 $(SC_x, SC_y, SC_z, 1)$ 、 $(SL_x, SL_y, SL_z, 1)$ 、 $(SR_x, SR_y, SR_z, 1)$ 、 $(SP_x, SP_y, SP_z, 1)$, 肩部中间骨骼点、左肩骨骼点和右肩骨骼点决定的平面方程如公式 (a) 所示,肩部中心骨骼点、脊柱中心骨骼点决定的直线方程如公式 (b) 所示:

$$A_p x + B_p y + C_p z + D_p = 0 \quad (a)$$

其中,平面方程系数 A_p 、 B_p 、 C_p 、 D_p 可由共面的肩部中间骨骼点、左肩骨骼点和右肩骨骼点在 Kinect 坐标系中的坐标通过待定系数法计算得到;

$$\frac{x - SC_x}{SP_x - SC_x} = \frac{y - SC_y}{SP_y - SC_y} = \frac{z - SC_z}{SP_z - SC_z} \quad (b)$$

由几何知识可以得到头部骨骼点坐标 $(H_x, H_y, H_z, 1)$ 与前面所述平面、直线的距离分别为 D_{plane} 、 D_{line} , 阈值 TH_{plane} 表示头部骨骼点与肩部中间骨骼点、左肩骨骼点、右肩骨骼点决定的平面的距离最大值, 阈值 TH_{line} 表示头部骨骼点与肩部中心骨骼点、脊柱中心骨骼点决定的直线的距离最大值, TH_{plane} 、 TH_{line} 可根据精度要求自行设定;若满足关系式 (c), 则得到了当前头部骨骼点在 Kinect 坐标系中的齐次坐标, 否则继续利用 Kinect 进行骨骼点的搜索;

$$D_{\text{plane}} \leq TH_{\text{plane}} \text{ 且 } D_{\text{line}} \leq TH_{\text{line}} \quad (\text{c}).$$

5. 根据权利要求 1 所述的基于 Kinect 的自动面部追踪方法,其特征在于:所述步骤 (3) 中计算当前摄像机的外参数和当前 Kinect 的外参数的方法为:如果选择第一种位置放置方式,则直接将步骤 (1) 中摄像机的外参数和 Kinect 的外参数分别作为当前摄像机的外参数和当前 Kinect 的外参数,如果选择第二种位置放置方式,则需要根据 Kinect 相对于上一位置转动的俯仰角度对 Kinect 的外参数进行修正,然后再将步骤 (1) 中摄像机的外参数和修正后的 Kinect 的外参数分别作为当前摄像机的外参数和当前 Kinect 的外参数。

6. 根据权利要求 1 所述的基于 Kinect 的自动面部追踪方法,其特征在于:所述步骤 (4) 中利用摄像机坐标系中的头部骨骼点坐标获取摄像机在水平和垂直方向需要旋转的角度和摄像机需要缩放的倍数的方法为:摄像机在齐次坐标系下的方向向量为 $\vec{dir} = (0, 0, 1, 0)$,头部骨骼点在摄像机坐标系中的齐次坐标为 $H = (X_{op}, Y_{op}, Z_{op}, 1)$,然后将 H 向 xOy 面投影得到投影点 P ,其坐标为 $(X_{op}, 0, Z_{op}, 1)$,计算两个向量, $\vec{OH} = (X_{op}, Y_{op}, Z_{op}, 0)$, $\vec{OP} = (X_{op}, 0, Z_{op}, 0)$,水平方向的旋转角度 θ_{pan} 和垂直方向的旋转角度 θ_{tilt} 分别由公式 (d) 和 (e) 得到:

$$\theta_{\text{pan}} = \arccos \frac{\vec{OP} \cdot \vec{dir}}{|\vec{OP}| |\vec{dir}|} \quad (\text{d})$$

$$\theta_{\text{tilt}} = \arccos \frac{\vec{OP} \cdot \vec{OH}}{|\vec{OP}| |\vec{OH}|} \quad (\text{e});$$

利用公式 (f) 计算头部骨骼点坐标与摄像机的距离,记作 dis ,设摄像机未缩放时,即初始状态,面部与摄像机的距离为 d_0 ,则缩放倍数 $Zoom$ 由公式 (g) 计算出:

$$dis = \sqrt{X_{op}^2 + Y_{op}^2 + Z_{op}^2} \quad (\text{f})$$

$$Zoom = \frac{dis}{d_0} \quad (\text{g}).$$

一种基于 Kinect 的自动面部追踪方法

技术领域

[0001] 本发明涉及一种自动面部追踪方法,特别涉及一种基于 Kinect 的自动面部追踪方法,属于人机交互领域。

背景技术

[0002] 人机交互技术是指通过计算机的输入输出设备,以便于人们使用的方式实现人与计算机之间有效交流的技术。面部追踪技术是人机交互领域的一项重要技术,它在视频会议、面部表情细节获取、消费电子类产品、面部运动捕捉以及自然人机交互环境中具有广泛的应用前景。

[0003] 现有的自动面部追踪方法通常是基于二维图像的,该类方法不仅涉及到一般的图像处理,而且可能设计到图像的理解和描述,如模式识别、人工智能等,相对复杂。其存在的主要问题在于对图像的分辨率有一定的要求,且该类方法对周围的光照环境敏感,光照条件不佳可能会影响追踪的准确性;人体面部的头发、饰物等遮挡物会减少人脸的部分局部特征,会对面部的追踪造成影响,甚至无法追踪,虽然可以通过人工智能进行补偿,但较为复杂且对实时性有影响;交互应用方面,该类方法只利用了二维的图像信息,无法鲁棒地处理人体部位在深度方向上的交互动作。

[0004] 微软公司的自然交互设备 Kinect,能够实时地捕捉人体 20 个骨骼点的三维空间位置,这对于人体追踪很有帮助。Kinect 采用主动红外线技术进行深度探测,在识别人体骨骼的同时,避免了周围光照条件和遮挡的影响。可编程旋转摄像机能够接收编程命令,按照设定的角度在水平和垂直方向上旋转,摄像机可通过编程命令来实现按需变焦。

[0005] Kinect 可以实现较为准确的人体骨骼点追踪,可编程旋转摄像机能够实现对于给定位置的图像或视频的捕获。因此,本发明的基于 Kinect 的自动面部追踪方法将二者进行结合,利用 Kinect 获取三维骨骼点数据,然后利用该数据驱动摄像机旋转,实现对面部的自动追踪。

发明内容

[0006] 本发明的目的是:结合可提取骨骼点三维信息的 Kinect 和可编程旋转摄像机,解决基于图像的自动面部追踪所面临的光照环境问题以及在交互应用中的不足,提供一种自动面部追踪方法,该方法在不同光照条件下和交互应用中,可以有效地利用 Kinect 提供的骨骼点三维信息,提高自动面部追踪的鲁棒性和准确性。

[0007] 为了实现上述目的,本发明所采取的技术方案是:本方法通过将 Kinect 与可编程旋转摄像机标定到同一个坐标空间中,对 Kinect 获取的人体骨骼数据序列进行处理和提取,来驱动摄像机对人的面部进行自动追踪,具体包含的步骤如下:

[0008] (1) 将追踪设备放置在稳定的位置,追踪设备由 Kinect 和可编程旋转摄像机组成,通过标定得到摄像机的内参数、Kinect 的内参数以及摄像机的外参数、Kinect 的外参数;

[0009] (2) 通过 Kinect 获取人体骨骼数据序列,并从中提取头部骨骼点在 Kinect 坐标系中的坐标,提取方法为直接提取,或者是在提取过程中结合骨骼点之间的约束关系进行优化;

[0010] (3) 完成步骤 (2) 以后,利用步骤 (1) 中得到的摄像机的外参数和 Kinect 的外参数,按照在步骤 (1) 中追踪设备位置放置方式,计算当前摄像机的外参数和当前 Kinect 的外参数并利用它们得到 Kinect 坐标系到摄像机坐标系的变换矩阵;

[0011] (4) 利用步骤 (3) 中获取的变换矩阵将步骤 (2) 中得到的头部骨骼点在 Kinect 坐标系中的坐标变换到摄像机坐标系中,得到头部骨骼点在摄像机坐标系中的坐标并利用该坐标计算摄像机在水平和垂直方向需要旋转的角度以及需要缩放的倍数;

[0012] (5) 根据步骤 (4) 中求得的角度和倍数,驱动摄像机进行相应的旋转和缩放,使得摄像机能够追踪人的面部运动。

[0013] 本发明是在将 Kinect 与摄像机标定到同一坐标空间的基础上,于 Kinect 的搜索范围内,连续采集人体骨骼数据序列并从中提取头部骨骼点的三维坐标,通过变换矩阵将其变换到摄像机坐标系中,根据该坐标计算摄像机需要旋转的角度和缩放倍数,从而完成自动的面部追踪。

[0014] 本发明的优点在于:

[0015] (1) 本发明利用了 Kinect 主动红外线技术获取的三维信息来驱动摄像机进行自动面部追踪,而红外线技术对于光照环境不敏感,甚至于没有环境光的情况下也能获取三维信息,有效克服了现有基于图像的自动面部追踪在周围光照不满足要求的情况下无法进行面部追踪的问题。

[0016] (2) 本发明充分利用人体骨骼点的三维信息来实现自动面部追踪,从而有助于交互应用方面,可以鲁棒地处理人体部位在深度方向上的交互动作,克服了现有基于图像的自动面部追踪的不足。

[0017] 本发明的应用前景包括:可应用于虚拟沉浸式系统中对于运动或静止的人体面部细节或面部表情的实时获取;可应用于动态或静态人脸建模中对于面部纹理的实时获取;可应用于自然交互式视频会议中,视频会议环境为三维虚拟场景,与会人员以人偶模型或真实人体模型为替身在虚拟会议场景中活动,同时显示所有与会人员(包括自己)的视频影像,该视频影像通过基于 Kinect 的自动面部追踪方法获得,与会人员在 Kinect 可搜索的范围内运动时,仍然可以自动捕获与会人员的面部,并将其显示在视频影像中。

附图说明

[0018] 图 1 是本发明中自动面部追踪方法的整体示意图;

[0019] 图 2 是本发明中自动面部追踪方法的流程图;

[0020] 图 3 是本发明中头部附近骨骼点示意图;

[0021] 图 4 是本发明中 Kinect 坐标系转换到摄像机坐标系的示意图;

[0022] 图 5 是本发明中摄像机旋转角计算的示意图;

[0023] 图 6 是实际硬件环境;

[0024] 图 7 是站立姿势效果图;

[0025] 图 8 是蹲姿势效果图。

具体实施方式

[0026] 下面结合附图和实施示例对本发明作进一步详细地描述：

[0027] 本发明中自动面部追踪方法的典型实施环境如图 1 所示，Kinect 和可编程旋转摄像机连接至计算机，人体位于 Kinect 的可搜索范围内，即可开始进行自动面部追踪。

[0028] 本发明基于 Kinect 的自动面部追踪方法的流程如图 2 所示，其步骤如下：

[0029] (1) 首先将追踪设备放置在稳定的位置，追踪设备由 Kinect 和可编程旋转摄像机组成，Kinect 和可编程旋转摄像机可以有两种位置放置方式，第一种是在 Kinect 上集成或固定摄像机，第二种是将 Kinect 和摄像机分别独立放在不同的位置，采用任意一种方法放置后，通过标定得到摄像机的内参数、Kinect 的内参数以及摄像机的外参数、Kinect 的外参数；

[0030] (2) 从 Kinect 的人体骨骼数据序列中直接提取的头部骨骼点坐标可以直接使用，也可利用骨骼点之间的约束关系来进行优化。若采用优化方法，如图 3 所示，头部骨骼点附近的骨骼点包括肩部中心骨骼点（或说颈部骨骼点）、左肩骨骼点、右肩骨骼点和脊柱中心骨骼点，它们之间存在着一定的约束关系，肩部中间骨骼点、左肩骨骼点和右肩骨骼点可认为是共面，肩部中心骨骼点、脊柱中心骨骼点可认为共线，头部骨骼点受到这二者的约束，虽然头部可以绕轴转动，但活动范围很有限，这意味着其距离肩部中间骨骼点、左肩骨骼点和右肩骨骼点决定的平面以及肩部中心骨骼点、脊柱中心骨骼点决定的直线很近。假设的头部骨骼点、肩部中心骨骼点、左肩骨骼点、右肩骨骼点和脊柱中心骨骼点的坐标分别为 $(H_x, H_y, H_z, 1)$ 、 $(SC_x, SC_y, SC_z, 1)$ 、 $(SL_x, SL_y, SL_z, 1)$ 、 $(SR_x, SR_y, SR_z, 1)$ 、 $(SP_x, SP_y, SP_z, 1)$ ，肩部中间骨骼点、左肩骨骼点和右肩骨骼点决定的平面方程如公式 (1) 所示，肩部中心骨骼点、脊柱中心骨骼点决定的直线方程如公式 (2) 所示。

$$[0031] \quad A_p x + B_p y + C_p z + D_p = 0 \quad (1)$$

[0032] 其中，平面方程系数 A_p 、 B_p 、 C_p 、 D_p 可由共面的肩部中间骨骼点、左肩骨骼点和右肩骨骼点在 Kinect 坐标系中的坐标通过待定系数法计算得到；

$$[0033] \quad \frac{x-SC_x}{SP_x-SC_x} = \frac{y-SC_y}{SP_y-SC_y} = \frac{z-SC_z}{SP_z-SC_z} \quad (2)$$

[0034] 由几何知识可以得到头部骨骼点 $(H_x, H_y, H_z, 1)$ 与平面、直线的距离分别为 D_{plane} 、 D_{line} 。阈值 TH_{plane} 表示头部骨骼点与肩部中间骨骼点、左肩骨骼点、右肩骨骼点决定的平面的距离最大值， TH_{line} 表示头部骨骼点与肩部中心骨骼点、脊柱中心骨骼点决定的直线的距离最大值， TH_{plane} 、 TH_{line} 可根据精度要求自行设定。若满足关系式 (3)，则得到了当前头部骨骼点在 Kinect 坐标系中的齐次坐标，否则继续利用 Kinect 进行骨骼点的搜索。

$$[0035] \quad D_{plane} \leq TH_{plane} \text{ 且 } D_{line} \leq TH_{line} \quad (3)$$

[0036] (3) 利用步骤 (1) 中得到的摄像机的外参数和 Kinect 的外参数，按照在步骤 (1) 中选择的 Kinect 和可编程旋转摄像机的位置放置方式，计算当前摄像机的外参数和当前 Kinect 的外参数并利用它们得到 Kinect 坐标系到摄像机坐标系的变换矩阵，方法如下：如果选择在 Kinect 上集成或固定摄像机，则可直接将步骤 (1) 中摄像机的外参数和 Kinect 的外参数分别作为当前摄像机的外参数和当前 Kinect 的外参数，如果步骤 (1) 中选择将 Kinect 和摄像机分别独立放在不同的位置，则需要根据 Kinect 相对于上一位置转动的

俯仰角度对 Kinect 的外参数进行修正,然后再将步骤 (1) 中摄像机的外参数和修正后的 Kinect 的外参数分别作为当前摄像机的外参数和当前 Kinect 的外参数;

[0037] (4) 利用步骤 (3) 中获取的变换矩阵将步骤 (2) 中得到的头部骨骼点在 Kinect 坐标系中的坐标变换到摄像机坐标系中,图 4 展示了从 Kinect 坐标系转换到摄像机坐标系的过程,其中 E_s 为当前摄像机外参数得到的外参矩阵, E_k^{-1} 为当前 Kinect 外参数得到的外参矩阵的逆矩阵, $E_s E_k^{-1}$ 即为变换矩阵;而后需要转换为摄像机能够识别的数据格式,即在水平方向和垂直方向上需要旋转的角度以及用于确定缩放倍数的头部骨骼点与摄像机坐标系原点的距离。如图 5 所示,摄像机在齐次坐标系下的方向向量为 $\vec{dir} = (0, 0, 1, 0)$,头部骨骼点在摄像机中的坐标为 $H = (X_{op}, Y_{op}, Z_{op}, 1)$,将其向 xOy 面投影得到投影点为 $P = (X_{op}, 0, Z_{op}, 1)$,计算两个向量, $\vec{OH} = (X_{op}, Y_{op}, Z_{op}, 0)$, $\vec{OP} = (X_{op}, 0, Z_{op}, 0)$,水平方向的旋转角度 θ_{pan} 和垂直方向的旋转角度 θ_{tilt} 分别由公式 (4) 和 (5) 得到:

$$[0038] \quad \theta_{pan} = \arccos \frac{\vec{OP} \cdot \vec{dir}}{|\vec{OP}| |\vec{dir}|} \quad (4)$$

$$[0039] \quad \theta_{tilt} = \arccos \frac{\vec{OP} \cdot \vec{OH}}{|\vec{OP}| |\vec{OH}|} \quad (5)$$

[0040] 利用公式 (6) 计算头部骨骼点坐标与摄像机的距离,记作 dis ,设摄像机未缩放时(即初始状态)面部与摄像机的距离为 d_0 ,则缩放倍数 $Zoom$ 可由公式 (7) 计算出。

$$[0041] \quad dis = \sqrt{X_{op}^2 + Y_{op}^2 + Z_{op}^2} \quad (6)$$

$$[0042] \quad Zoom = \frac{dis}{d_0} \quad (7)$$

[0043] (5) 完成步骤 (4) 以后,需要将得到的摄像机在水平方向和垂直方向的旋转角度、缩放倍数通过编程命令发送给可编程旋转摄像机,驱动摄像机进行面部追踪。

[0044] 本发明可采用微软提供的 Kinect SDK 对 Kinect 进行操作,将计算机与可编程旋转摄像机相连,控制摄像机和实现通信。实际硬件环境如图 6 所示,站立效果如图 7 所示,半蹲效果如图 8 所示,图中矩形框分别标识出设备、真实人体、面部追踪结果(显示在屏幕上)。

[0045] 本发明中未详细阐述的部分属于本领域技术人员的公知技术。

[0046] 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,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

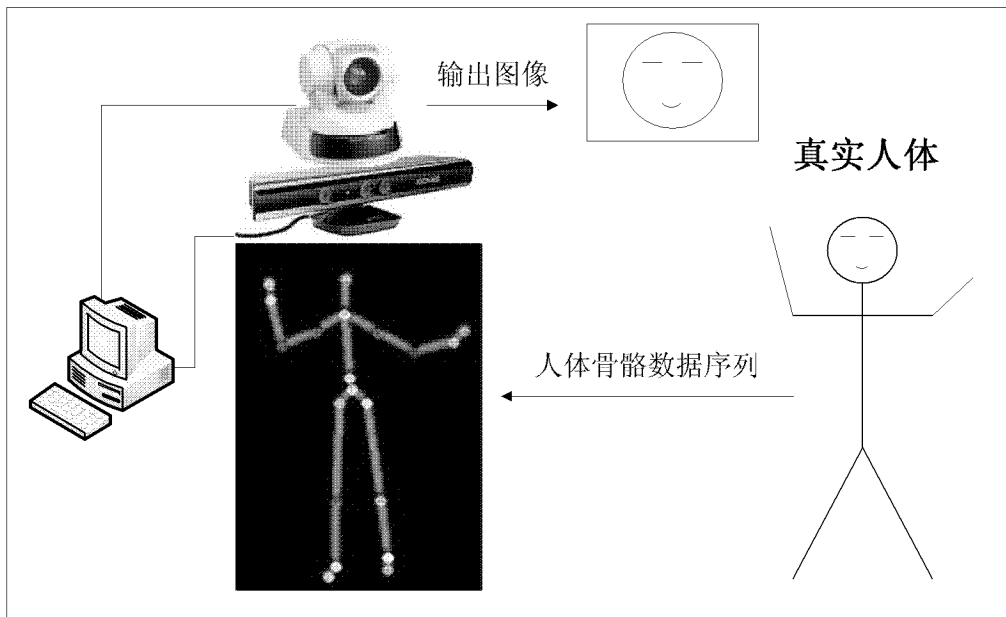


图 1

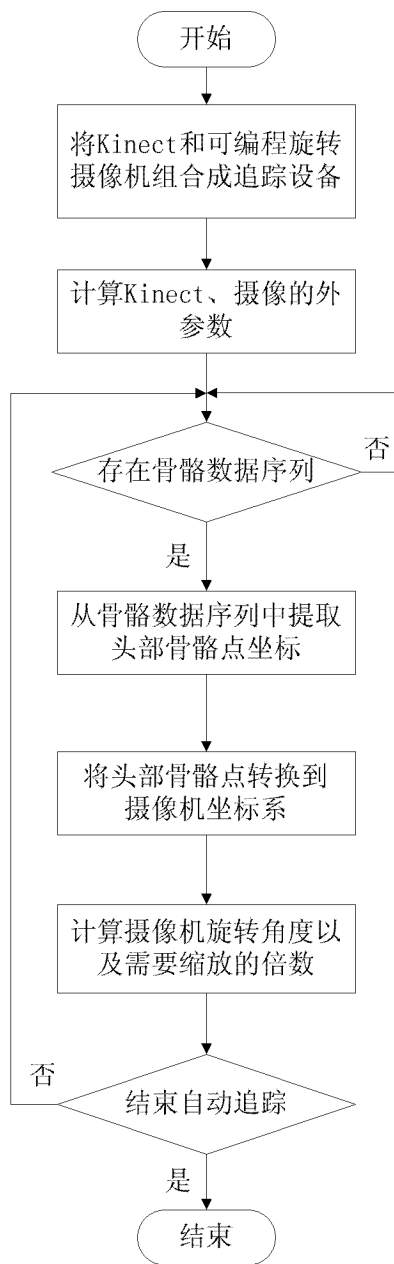


图 2

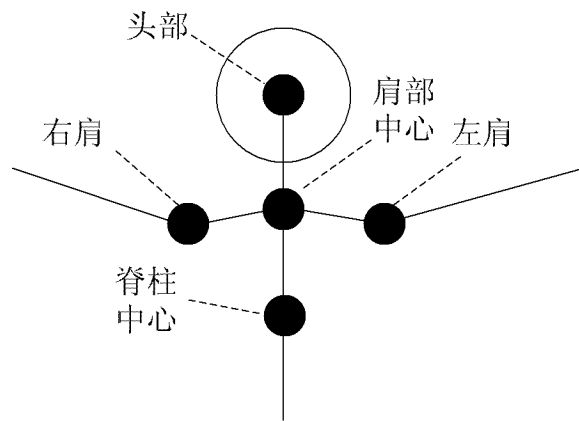


图 3

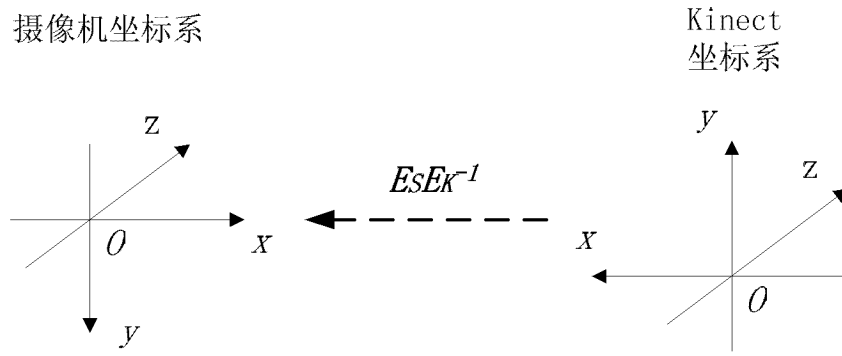


图 4

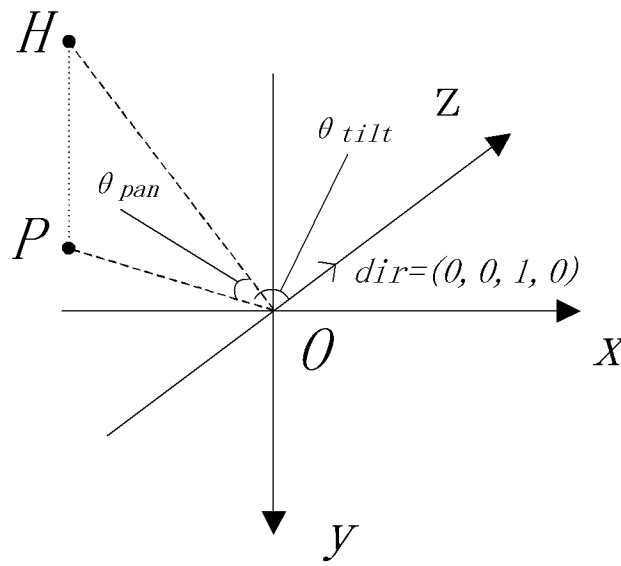


图 5



图 6

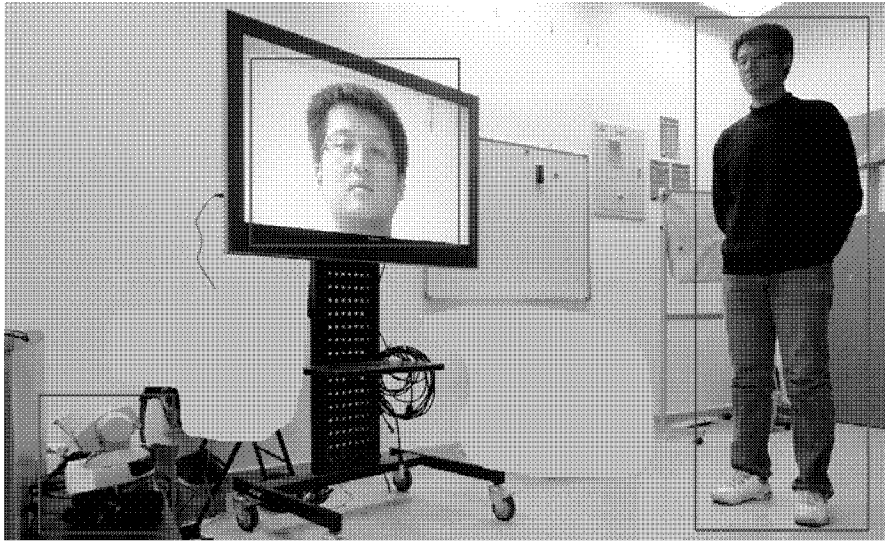


图 7



图 8