

一种具有逼真效果的虚拟人动画生成方法^{*}

雷涛^{1,2}, 罗维薇¹, 王毅²

(1. 兰州交通大学 电子与信息工程学院, 兰州 730070; 2. 西北工业大学 电子信息学院, 西安 710072)

摘要: 针对传统的虚拟人动画逼真度低、动画生成复杂、且动画存在飘移现象的问题,提出了一种基于视频关键帧获取运动数据的虚拟人动画生成方法。首先利用线性混合蒙皮算法建立了新的人体几何模型;其次利用正向运动学方法驱动虚拟人的骨骼模型,并利用平移补偿原则消除虚拟人运动中的飘移现象;最终实现了基于双目正交视觉获取关节运动参数的虚拟人动画。实验结果表明,新方法能逼真地模拟真实人体动作,且方法易于实现、实用性强。

关键词: 虚拟人; 正向运动学; 线性混合蒙皮; 平移补偿

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2012)10-3983-04

doi:10.3969/j.issn.1001-3695.2012.10.102

Method of realism emotion animation generation for virtual human

LEI Tao^{1,2}, LUO Wei-wei¹, WANG Yi²

(1. School of Electronic & Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China; 2. School of Electronics & Information, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Traditional virtual human animation has some drawbacks include low realism, complex generation process, and drift phenomenon, this paper proposed a novel generation approach of virtual human animation based on the video key frame obtained. First, it used linear blending skin algorithm to construct a new body geometry model, and then used forward kinematics method to drive the virtual human skeleton model. Furthermore, it introduced the principle of translation compensation, and it was able to eliminate drift phenomena in the virtual human movement. Finally, it carried out the virtual human animation according to the joint parameters of real human movement, which could be obtained by using binocular orthogonal video. The experimental results show that the new method can provide realistic virtual human animation for real human movement, and the method is easy to realize with strong practical applicability.

Key words: virtual human; forward kinematics; linear blending skin; translation compensation

0 引言

在虚拟环境中,利用虚拟人来模拟真实人的行为并对其进行评估是近年来信息科学领域的一个研究热点,它被广泛应用于动画、机械工程、医学、军事和空间探索等领域。目前,众多的科研机构及公司已经投入到虚拟人几何建模及动画生成的研究工作中,具有代表性的是蒙特利尔大学 Thalmann 建立的虚拟现实研究中心,宾夕法尼亚大学 Badler 建立的人体建模与仿真中心,加利福尼亚大学的 San Diego 等人带领的研究小组,华盛顿大学 Grochow 等人建立的虚拟人研究中心以及克莱斯勒汽车公司等。由于虚拟人动画生成的最终目的是逼真地模拟真实人体的动作及情感,因此,近年来研究人员提出了很多关于逼真动画的实现技术,这些新技术极大地推动了虚拟现实技术在军事和娱乐领域的快速发展。虚拟人动画生成主要包含面部动画^[1-3]和肢体动画。由于肢体动画是基于骨骼动画模型实现的,因此如何获取真实人体运动中各关节角度参数成为研究难点和热点。目前已有的技术大多是利用视频跟踪原理实现的^[4-6]。例如,针对视频中的运动人体动作跟踪,

Cheung 等人^[4]提出了 SFS 算法(shape-from-silhouette),该算法可以用于人体建模和带标记的运动跟踪,且跟踪精度较高;缺陷是需要对人体进行标记,人体运动容易受环境影响,且遮挡关节的数据难以获取。Urtasuna 等人^[5]针对蒙特卡罗算法计算量较大的问题,提出了一种针对运动跟踪进行学习的新模型,并利用迭代优化的方法实现精确跟踪。Zou 等人^[7]针对无标记的运动人体进行跟踪,并成功实现了跟踪数据在虚拟模型上的运动匹配。这些方法均简化了虚拟人动画实现过程中获取运动关节参数的问题。

就现有的骨骼动画原理而言,常用的方法是正向运动学和逆向运动学,而对皮肤网格运动的控制主要依赖于线性混合蒙皮算法,对 3D 人体几何模型的蒙皮过程主要依赖动画师的经验手动实现。因此 Li 等人^[8]提出了一种自动蒙皮算法,从而解决了手动蒙皮处理中依赖于人的问题;缺陷在于对某些关节的蒙皮效果并不理想。Chen 等人^[9]就同样的问题,提出了一种更为高效的自动蒙皮算法,该算法仅利用少量的人工交互就可实现较好的蒙皮结果。此外,还有多种方法可在虚拟人动画实现过程中减少人机交互并提升虚拟人动画的逼真性。然而,对于现有的商业人体几何模型,要实现其骨骼动画,只有小关

收稿日期: 2012-03-02; **修回日期:** 2012-04-06 **基金项目:** 国家自然科学基金资助项目(60903127,61202314);甘肃省青年科技研究基金资助项目(2010GS04061)

作者简介: 雷涛(1981-),男,陕西大荔人,副教授,博士(后),主要研究方向为模式识别、虚拟现实、图像处理等(leitaoly@163.com);罗维薇(1977-),女,讲师,硕士,主要研究方向为图像处理、虚拟现实等;王毅(1979-),女,副教授,博士,主要研究方向为图像处理、虚拟现实等。

节运动尚可实现。对于较为复杂的动作和表现而言,通常涉及到平移和旋转的混合运算,而导出的网格模型通常以骨盆为根节点,这就限制了虚拟人骨盆以下的腿部关节运动,实现的动画具有飘移感。为了解决该问题,根据目前已有的虚拟人几何建模和运动控制技术,本文构建了具有高逼真效果的虚拟人体几何模型,并提出了一种新的平移补偿原则,从而消除了虚拟人动画中的飘移现象。在此基础上选择了基于视频关键帧技术的人体运动关节参数获取方法,最后基于 DirectX 和 VC2005 环境,实现了具有逼真效果的虚拟人体动画。

1 虚拟人几何建模

虚拟人建模过程可分为表面网格模型获取、创建骨骼、蒙皮和模型导出四个步骤。

1.1 表面网格模型获取

人体曲面数据的获取是建立人体模型的重要组成部分,人体形状为复杂的曲面,要对其进行较为精确的测量并获得全面细致的人体数据是很困难的。本文采用了 e-frontier 公司 Poser7 中的人体模型,将人体模型转换成 3DSmax 支持的 3ds 文件后,再用 Right Hemisphere 公司的 Deep exploration 5 为模型添加纹理信息,从而得到所需要的人体几何网格模型。

1.2 创建骨骼

虽然真实人体有 206 块骨骼,但真正能表现人体运动的骨骼数目并不多,因此在虚拟人运动仿真过程中,没有必要将每块骨骼都定义出来。此外,为了在 PC 平台上实现实时的虚拟人运动,必须对人体的骨骼结构进行一定程度的简化。这样就可以大幅度降低人体骨骼结构的复杂性,实现虚拟人的实时运动。本文定义的虚拟人包含 56 块骨骼,其层次结构如图 1 所示。

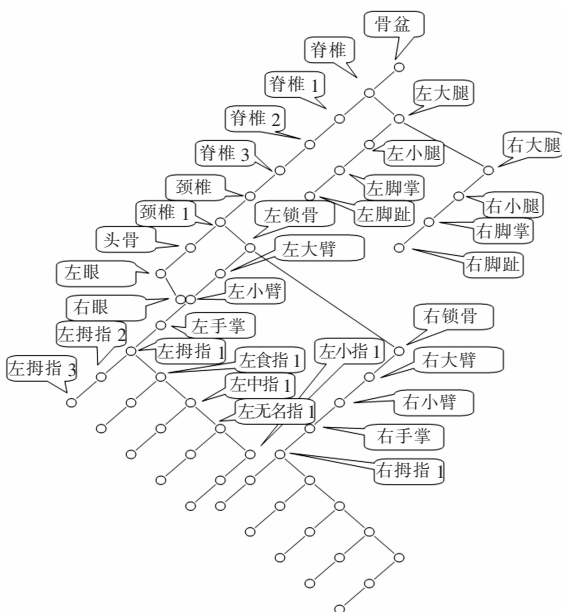


图 1 骨骼层次连接结构

1.3 蒙皮

在骨骼动画网格模型中,网格模型是由大量顶点构成的,如果任意一个顶点仅受一块骨骼影响,那么网格模型在运动时很容易出现裂缝,尤其是结合部位,如肩关节、肘关节等。为了解决该问题,研究人员提出了线性混合蒙皮算法。线性混合蒙

皮算法是将静态的网格和内在的骨骼框架结合在一起,网格上的每个顶点通过蒙皮权重将骨骼空间变换映射到网格顶点,从而实现皮肤层的运动^[10]。在皮肤层中,每个顶点都对应两个参数,一个控制它的骨骼,另一个表示控制程度的权重,据此就可实现从骨骼层到皮肤层的映射。皮肤层顶点的变形和运动通过变换影响该顶点所有骨骼来实现。变形后的皮肤层网格顶点的位置可以用式(1)来计算:

$$\bar{V} = \sum_{i=1}^n \omega_i \times M_{i,k} \times M_{i,0}^{-1} \times V_0 \quad (1)$$

其中: ω_i 表示第 i 块骨骼对网格顶点控制的权重,为使网格在运动时不出现开裂, ω_i 应满足约束

$$\sum_{i=1}^n \omega_i = 1 \quad (2)$$

V_0 表示在世界坐标系中要发生运动的顶点的初始位置; $M_{i,0}$ 表示在初始姿态下从第 i 块骨骼的局部坐标系到世界坐标系的变换矩阵; $M_{i,0}^{-1}$ 则表示在初始姿态下该骨骼从世界坐标系到该骨骼局部坐标系的变换矩阵; $M_{i,k}^{-1} \times V_0$ 为在第 i 块骨骼的局部坐标系中表示的顶点初始位置; $M_{i,k}$ 为世界坐标系中表示的运动变换矩阵; $M_{i,k} \times M_{i,0}^{-1} \times V_0$ 表示在蒙皮权重为 1 的情况下,第 i 块骨骼的局部坐标系内该骨骼所影响的顶点从初始位置到运动后位置的变换。

以肘部的点为例,如图 2 所示。 P 和其附近的点受到大臂和小臂两段骨骼的影响,小臂(红色部分)由初始伸直状态转动到弯曲的状态。初始状态时, P 点位于 P'' 的位置,如果 P 点仅仅受到小臂的骨骼影响,则在曲臂的动作中, P 点会移动到 P' 点;如果 P 点仅受到大臂的骨骼影响,曲臂时, P 点待在于 P'' 点的位置。正是由于 P 同时受到大臂和小臂两段骨骼的控制(此处假设各自权重都为 0.5), P 正好移动到大臂和小臂的中心位置, P 周围的点也由于受到两块骨骼的影响,产生了逐渐变形的效果。

通过线性混合蒙皮算法,将骨骼运动和皮肤层的网格运动相分离,只需研究骨骼运动便可利用蒙皮权重这一映射关系实现网格的运动。

1.4 模型导出

模型建好之后,必须导出 .X 格式才能在 D3D 开发的程序中调用,而 3DMax 中的 panda DirectX Export Plugin 插件可用来导出 .X 文件。最终完成的虚拟人几何模型如图 3 所示。

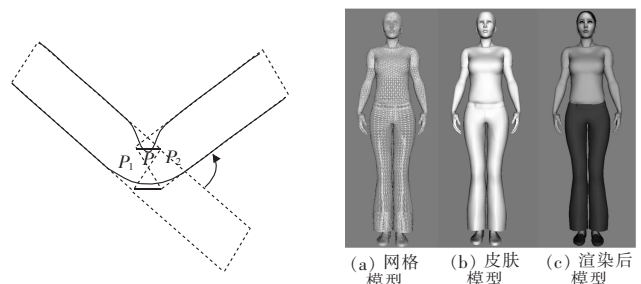


图 2 曲臂动作下网格顶点运动示意图

图 3 虚拟人几何模型

2 基于正向运动学的虚拟人运动控制

骨骼动画是目前流行的也是最复杂的角色动画,其原理源于自然界大多数动物的真实运动行为。由于动物身体的肌肉和皮肤均依附于骨骼,当骨骼开始运动时,依附于对应骨骼的肌肉和皮肤都随骨骼一起运动。用数学模型表示骨骼运动的

原理,则骨骼一般对应两个矩阵:初始变换矩阵,表示骨骼的初始位姿;组合变换矩阵,用于计算骨骼的各种运动,从而实现角色动画。根据骨骼运动原理,可以对角色模型进行逐帧渲染,以实现角色动画,从而需要更新骨骼层次结构,组合每个连续的变换,将上层骨骼运动传递到下层骨骼中^[11,12]。

正向运动学通过给定状态向量来确定各关节的位置,当人体模型姿势发生变化时,正向运动学的解决过程是:依照“根节点(骨盆)→脊椎节点→锁骨关节点→肩关节点→肘关节点→腕关节点”这样一个过程“正向”地改变各关节点(身体段)的位置和方位,最后确定这一开放链末端(手部、手指尖)在空间的位置和方位。如图 4 所示,在一个三关节的胳膊骨骼系统中,下一级骨骼的位置要受到其所有上级骨骼状态的影响。如果绕着肩部关节点旋转大臂,则小臂和手必然跟随肩部做相应的旋转;同理,如果旋转小臂,手也跟着做相应的旋转。一块骨骼的位姿首先受到该块骨骼的组合旋转矩阵影响,此外还要受该骨骼的所有父骨骼的组合旋转矩阵影响。

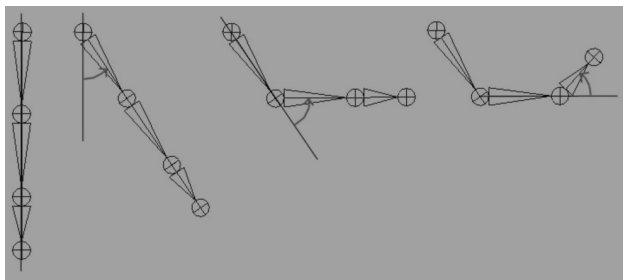


图 4 胳膊旋转示意图

在骨骼模型中,骨骼子节点坐标系与父节点坐标系之间的关系可以由矩阵确定,如式(3)所示。变换可以分为平移、旋转以及放缩变换,而这里对骨骼动画的实现只涉及到平移和旋转变换,根据骨骼传递原理,只需知道平移矩阵和组合旋转矩阵即可。这里的组合旋转矩阵包含 $i-1$ 个矩阵,对于第 i 级骨骼而言,它与 $i-1$ 级骨骼之间的关系为

$$M = T(t_x, t_y, t_z) \times R_z(\lambda) \times R_y(\beta) \times R_x(\alpha) \quad (3)$$

其中:

$$T(t_x, t_y, t_z) = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

其中: T 表示骨骼根节点的局部坐标系相对于世界坐标系的变换矩阵,为平移矩阵。

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) & 0 \\ 0 & \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$R_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$R_z(\gamma) = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

$R_x(\alpha)$ 、 $R_y(\beta)$ 和 $R_z(\gamma)$ 分别表示绕 X 、 Y 和 Z 轴的旋转矩阵; $R_{x,y,z} = R_x(\alpha) \times R_y(\beta) \times R_z(\gamma)$ 表示总的旋转矩阵。根据式

(3), M 表示子节点在父节点局部坐标系内的相对位姿,为了将该节点在世界坐标系内渲染,则需要逆向递推计算,以得到该节点相对根节点的局部坐标位姿。若定义第 i 级骨骼是第 $i-1$ 级骨骼的子节点,则第 i 级骨骼的局部坐标系相对于第 $i-1$ 级骨骼的局部坐标系的变换矩阵为 ${}^{i-1}_i R$,由于多级骨骼的联合运动,从而第 i 级骨骼在世界坐标系中的变换矩阵为

$${}^w_i M = T(t_x, t_y, t_z) \times \left[\prod_{i=2, h=1}^{h=i-1} {}^{i-1}_i R(\alpha, \beta, \lambda) \right] \times {}^i M(x, y, z) \quad (8)$$

由式(8)可以看出,某一级骨骼的运动只对其下一级以后的骨骼运动产生影响,对该骨骼同一级和上一级的骨骼不会起作用。由此,可以通过运动链表示每个点的3D坐标与各个旋转角的关系,各点在世界坐标系中的位置可用同样方法求得。

3 平移补偿

利用 Poser 导出的人体几何模型在 3DMax 中进行蒙皮后,存在一个问题,即人体几何模型的世界坐标系在根关节处,而这是本文想避免的一个问题。由于人体的静态动作是全身运动,是以脚部为世界坐标系的,而这种人体几何模型导致的结果就是根关节以下的腿部均不能运动,如果运动,就会存在飘移现象。当根关节旋转时,整个身体就随着该关节的旋转而旋转,从而产生飘移现象。为了解决该问题,从几何角度考虑,可以得到人体下肢运动的原理,如图 5 所示。

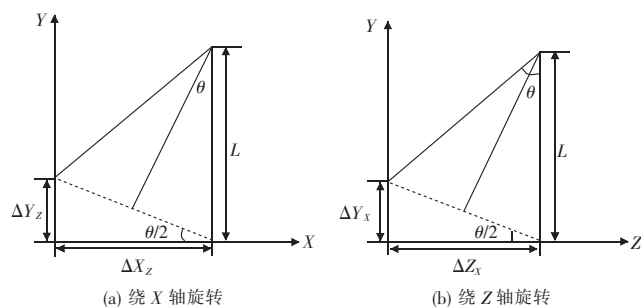


图 5 下肢旋转

定义人体腿部的长度为 L ,如果根关节发生了偏移而腿部相对根关节不动时,则如图 5 所示。由根关节的旋转推导出根关节的平移补偿数据,利用根关节的平移补偿可以固定脚部不动,这样便可以得到更为逼真的人体动画,而且计算复杂度也较小。

绕 X 轴旋转:

$$\begin{cases} \Delta Z_x = L \times \sin \theta \\ \Delta Y_x = 2L \times \left(\sin \frac{\theta}{2} \right)^2 \end{cases} \quad (9)$$

绕 Z 轴旋转:

$$\begin{cases} \Delta X_z = L \times \sin \theta \\ \Delta Y_z = 2L \times \left(\sin \frac{\theta}{2} \right)^2 \end{cases} \quad (10)$$

因此,当盆骨进行旋转时,根据角度值计算其平移距离,并赋给盆骨的平移矩阵。该方法一定程度上解决了动画逼真度不高的问题。

4 具有逼真效果的虚拟人动画实现

为了实现具有真实感的虚拟人动画,本文利用关键帧方法

获取运动人体参数。即首先录制真实的运动人体视频,然后提取关键帧,利用双目正交视频中人体关节角度测量的方法得到各个关节的角度参数;然后利用骨骼动画原理实现3D虚拟人肢体动画。利用 VisualC++ 2005 联合 DirectX10.0 开发虚拟人肢体动画程序,其程序流程如图6所示。

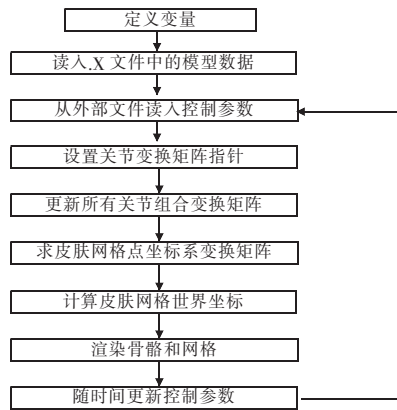


图6 虚拟人肢体动画程序流程

根据图6的流程,利用 VisualC++ 2005 联合 DirectX10.0 开发虚拟人肢体动画程序,其程序各步骤的描述如下:

a) 定义运动关节运动参数集合 $J_x = \{J_1^x, J_2^x, J_3^x, \dots, J_{51}^x\}$, $J_y = \{J_1^y, J_2^y, J_3^y, \dots, J_{51}^y\}$, $J_z = \{J_1^z, J_2^z, J_3^z, \dots, J_{51}^z\}$, 三个参数集合分别表示关节绕 X 轴、Y 轴和 Z 轴旋转的欧拉角集合。其中 $J_1^x = \{J_{11}^x, J_{12}^x, J_{13}^x, \dots, J_{151}^x\}$, $J_1^y = \{J_{11}^y, J_{12}^y, J_{13}^y, \dots, J_{151}^y\}$, $J_1^z = \{J_{11}^z, J_{12}^z, J_{13}^z, \dots, J_{151}^z\}$, $1 \leq i \leq 51$ 表示可运动的关节总共有 51 个, m 表示动画的帧数。

b) 建立关节运动参数与动画时间的函数关系。如果设动画的帧速度为 v fps, 则以时间 t 为自变量, 随时间的变化取不同的关节参数, 对于 J_{ij}^x, J_{ij}^y 和 J_{ij}^z 中 j 的取值为 $j = \text{round}(v \times t)$, v 可根据实际动画播放的需要选取, $24 \leq v \leq 30$, 本文取 $v = 25$ 。

c) 根据读入的关节运动参数, 计算组合变换矩阵。针对第 i 个关节在 j 时刻计算得到的组合变换矩阵为 $R_{ij} = R_{j_{ij}} \times R_{j_{ij}} \times R_{j_{ij}}$ 。由此可以计算出第 i 个关节在 j 时刻的初始矩阵更新为 $M_{ij}^{(i-j)} = R_{ij} \times M_{i(j-1)}$ 。该矩阵为局部坐标系内矩阵, 是第 i 个关节在第 $i-1$ 关节坐标系内的变换矩阵。最后利用世界矩阵得到关节在世界坐标系内的位姿。

d) 利用新时刻计算得到的初始矩阵按照骨骼遍历结构顺序进行递归计算, 使得父骨骼的旋转变换遍历到所有子骨骼中, 即 $M_{ij} = M_{ij} \times M_{j_1} \times \dots \times M_{j_n} \times M_{\text{world}}$, 从而计算出第 i 个关节在世界坐标系内的变换矩阵。

e) 将计算得到的最终变换矩阵与各骨骼块附带的皮肤网格的偏移矩阵相乘, 以计算得到变换后的偏移矩阵坐标系, 利用该矩阵可以渲染所有网格点到世界坐标系内: $M_i = M_{\text{offset}} \times M_{ij}$ 。

f) 利用式(1)计算依附于骨骼的皮肤网格顶点的运动(在世界坐标系内的位置)。

g) 进入下一时刻, 返回步骤 b) 循环执行, 直至时间结束。

5 实验结果及结论

本文基于 VisualC++ 2005 联合 DirectX10.0 开发虚拟人肢体动画程序, 计算机硬件采用 Intel Pentium® Dual E2200

CPU 2.2 GHz, 2 GB 内存, NVIDIA Ge-Force 9600GT 显卡。对实际录制的人体视频动作进行模拟, 录制的双目正交视频如图7所示。图8(a)实现的是虚拟人模拟真实人体动作(招手微笑)的动画结果, (b)实现的是揉眼哭泣的动画结果。最终实现的人体动画时长度为 8 s, 对视频动画进行抽样, 得到一些关键帧结果, 从不同的角度对视频进行截图, 每一行代表一个角度, 结果表明, 本文实现的虚拟人运动控制方法的确可以真实地模拟人体动作。

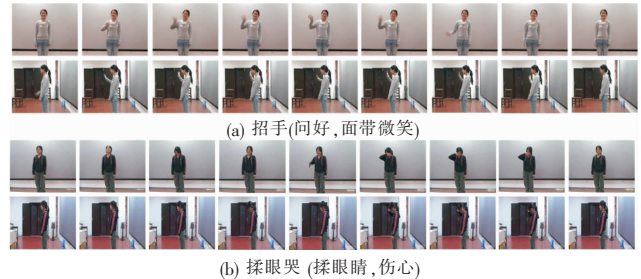


图7 实际录制的双目正交视频

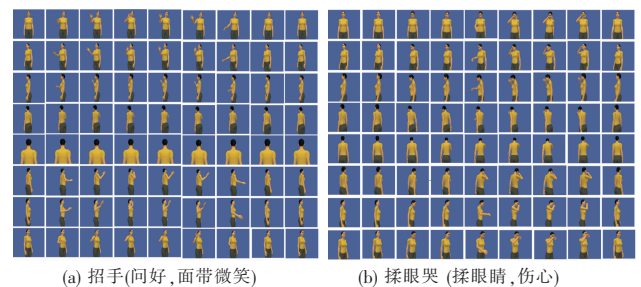


图8 虚拟人动画生成(从八个不同的角度观测虚拟人动画)

6 结束语

本文利用视频序列提取运动人体的关节信息, 再利用正向运动学对虚拟人肢体模型进行驱动, 取得了较好的效果。由于视频序列中存在着许多非刚性的和幅度比较小的运动, 因此在人工获取肢体运动参数时存在一定的误差, 然而这种误差在虚拟人重现真实人体动作时并不明显。显然, 本文提出的方法能较为逼真地模拟真实人体动作, 且方法简单、实用性强。

参考文献:

- [1] ZHANG Y, SIM T, TAN C L, et al. Anatomy-based face reconstruction for animation using multi-layer deformation [J]. *Journal of Visual Languages & Computing*, 2006, 17(2): 126-160.
- [2] VENKATARAMAN K, LODHA S, RAGHAVAN R. A kinematic-variational model for animating skin with wrinkles [J]. *Computers & Graphics*, 2005, 29(5): 756-770.
- [3] TANG Y M, HUI K C. The effect of tendons on foot skin deformation [J]. *Computer-Aided Design*, 2007, 39(7): 583-597.
- [4] CHEUNG K M, BAKER S, KANADE T. Shape-from-silhouette across time part II: applications to human modeling and markerless motion tracking [J]. *International Journal of Computer Vision*, 2005, 63(3): 225-245.
- [5] URTASUNA R, FLEETB D J, FUAA P. Temporal motion models for monocular and multiview 3D human body tracking [J]. *Computer Vision and Image Understanding*, 2006, 104(2-3): 157-177.
- [6] POPPE R. Vision-based human motion analysis: an overview [J]. *Computer Vision and Image Understanding*, 2007, 108(1-2): 4-18.

(下转第3992页)

(上接第 3986 页)

- [7] ZOU B J, CHEN S, CAO S. Automatic reconstruction of 3D human motion pose from uncalibrated monocular video sequences based on markerless human motion tracking[J]. **Pattern Recognition**, 2009, 42(7):1559-1571.
- [8] LI J T, LU G D. Skeleton driven animation based on implicit skinning [J]. **Computers & Graphics**, 2011, 35(5):945-954.
- [9] CHEN Cheng-hao, LIN I C, TSAI M H, *et al.* Lattice-based skinning and deformation for real-time skeleton-driven animation[C]//Proc of the 12th International Conference on Computer-Aided Design and Computer Graphics. [S. l.]:IEEE, 2011:306-312.
- [10] MOHR A, GLEICHER M. Building efficient, accurate character skins from examples[J]. **ACM Trans on Graphics**, 2003, 22(3):562-568.
- [11] MERRICK D, DWYER T. Skeletal animation for the exploration of graphs [C]//Proc of ACM International Conference on Information Visualisation. Darlinghurst: Australian Computer Society, 2004:61-70.
- [12] WU F C, MA W C, LIANG R H. Domain connected graph: the skeleton of a closed 3D shape for animation[J]. **The Visual Computer**, 2006, 22(2):117-135.