

doi: 10.3979/j.issn.1673-825X.2014.04.018

## 一种基于 Kinect 的角色骨骼动画方法

李红波,冉光勇,吴 渝,丁林建

(重庆邮电大学 计算机科学与技术学院网络智能研究所,重庆 400065)

**摘 要:** 针对当前骨骼动画制作方法存在实时性差、抗光照干扰能力弱、角色模型运动误差大的问题,提出了一种新的骨骼动画方法。采用微软 Kinect 设备获取人体三维骨骼坐标数据,用改进的 D-H(denavit-hartenberg)描述法在关节处建立虚拟坐标系,对捕捉对象单帧姿态进行抽象空间位形描述,再根据前后帧位形数据求解骨骼矩阵,对根节点进行位移调整形成骨骼动画。实验结果表明,该方法实时性与抗光照干扰能力强,角色动作与捕捉对象动作之间的相似度高。

**关键词:** 骨骼动画;实时;Kinect;骨骼数据

**中图分类号:** TP391.9

**文献标识码:** A

**文章编号:** 1673-825X(2014)04-0530-06

## A role skeletal animation method based on Kinect

LI Hongbo, RAN Guangyong, WU Yu, DING Linjian

(Institute of Web Intelligence, College of Computer Science and Technology,  
Chongqing University of Posts and Telecommunications, Chongqing 400065, P. R. China)

**Abstract:** Since the current skeletal animation production methods have problems of poor real-time and weak anti-interference ability for illumination, as well as big motion error in role model, a new skeletal animation method is proposed. Firstly, the three-dimensional coordinate data of human skeleton is obtained by using Microsoft's Kinect. Then an improved D-H(denavit-hartenberg) description method is used to establish the virtual coordinate system at the joints, and the single frame posture of the captured object is described by use of the configuration in the abstract space. Later, the bone matrix is solved based on the configuration data of front and rear frames. Finally, the displacement of the root node is adjusted to form the skeletal animation. Experimental results show that the proposed method has good real-time, anti-interference ability for illumination, and high similarity between the motion of the role and the captured object.

**Key words:** skeletal animation; real-time; Kinect; skeletal data

## 0 引 言

虚拟人骨骼动画是角色动画领域的主要研究内容<sup>[1]</sup>,其制作技术除人工调整关键帧形成骨骼动画外,还有基于视觉与运动捕获数据的方法。

基于视觉方法<sup>[2]</sup>的一般流程是:通过对包含人体运动序列的视觉素材进行跟踪分析,提取运动序列,接着对运动序列进行条件约束等处理,以此得到骨骼动画。该方法成本低、运动类型较丰富,但存在以下 3 方面不足:1) 利用二维数据生成三维运动数据存在信息缺失的问题<sup>[3]</sup>;2) 人体运动序列的获取过程易受光照强度影响,抗光照干扰能力弱<sup>[4]</sup>;3)

提前准备视觉素材的做法不能满足交互式环境对实时性的要求。

基于运动捕获数据的方法通常需要价格昂贵的运动捕捉设备,并且需要捕捉对象在关节处佩戴光、电学设备,致使人体运动灵活性低,动作获取成本高<sup>[5]</sup>。Park 等<sup>[6]</sup>提出了一种基于光学标记点的方法,密集的标记点集合提高了运动真实性,但降低了计算效率,无法满足实时性需求。Casas 等<sup>[7]</sup>通过使用 4D 运动图,将不同部分捕获数据进行分块匹配与连接,以此生成骨骼动画。该方法能够满足较多运动类型需求,但是容易出现裂缝现象。Boboc 等<sup>[8]</sup>提出了一种基于前向运动学的骨骼动画方法,

收稿日期:2013-03-03 修订日期:2014-05-30 通讯作者:李红波 lihongbo@cqupt.edu.cn

基金项目:“核高基”重大专项(2009ZX01038-002-002-2);科技部“原创动漫软件开发技术人才”计划扶持(2009-593)

**Foundation Items:** The “CHB” Major Project (2009ZX01038-002-002-2); The Program for Original Cartoon Software Programmer Supported by Ministry of Science and Technology of China(2009-593)

该方法以 Kinect 提取到的捕捉对象骨骼数据为基础,使用前向运动学求解角色模型关节点坐标数据,并对坐标数据进行平滑处理,使用处理后的数据调整模型,得到骨骼动画,此方法实时性强,但角色模型的运动与真实运动之间存在较大差别,同时,固定运动参数的做法降低了模型适用性。

针对上述问题,本文提出了一种新的解决方法,通过在 Kinect 骨骼模型的每个关节点处建立虚拟坐标系,对捕捉对象的单帧姿态进行抽象空间位形描述,利用位形数据求解关节点骨骼旋转矩阵,使用骨骼旋转矩阵对角色模型进行从端到根的关节方向与位置调节。实验结果表明,本文提出方法的实时性能与抗光照能力强,角色模型运动与真实运动相似度高。

## 1 相关技术基础

### 1.1 KINECT 虚拟人骨骼模型

Kinect 人体骨骼简化模型,如图 1 所示,包含了人体主要的 20 个关节节点,由 19 个骨骼段组成。按照人体构造结构,任意 2 个关节之间建立 Kinect 关节点父子关系,形成一个简化的最小人体骨骼模型。其中,臀部中心关节点是整个人体骨骼模型的根节点,它的三维坐标数据完全决定了整个人体的移动轨迹。臀部中心关节点有 6 个自由度,分别为 3 个平移自由度与 3 个旋转自由度,其他节点仅仅包括 3 个旋转自由度。

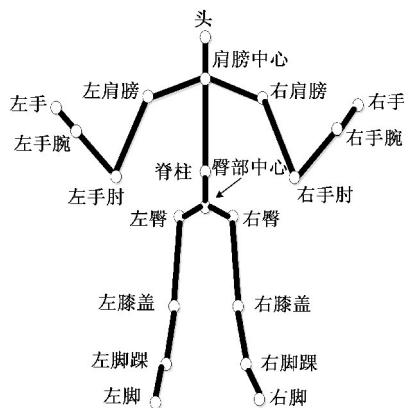


图 1 Kinect 骨骼模型

Fig. 1 Kinect skeleton model

### 1.2 骨骼动画基本原理

为了对骨骼动画原理进行描述,设  $p(j)$  代表关节点  $j$  的父节点,  $R(j)$  代表从父节点  $p(j)$  的局部坐标系到子节点  $j$  的局部坐标系的变化矩阵,则对于任意的骨骼关节点  $j$ ,从根节点局部坐标系转化到  $j$  的局部坐标系可以用矩阵  $F(j)$  表示<sup>[9]</sup>为

$$F(j) = R(0) \cdots R(p(j)) R(j), \quad (1)$$

当人体姿态发生变化时,我们用  $T(j)$  表示某一帧关节点  $j$  的骨骼旋转矩阵,则  $j$  在这一帧的变化矩阵为

$$F'(j) = F(p(j)) R(j) T(j), \quad (2)$$

假设网格节点  $v$  附着于骨骼关节点  $j$ ,  $v'$  为人体姿态发生变化后  $v$  在骨骼关节点  $j$  下的新位置,则新位置的计算公式为

$$v' = F'(j) F(j)^{-1} v, \quad (3)$$

在帧渲染时,通过(3)式便可求得皮肤顶点的新位置。

## 2 Kinect 骨骼动画方法

### 2.1 问题提出和总体思路

基于视觉生成骨骼动画的方法虽然与实际运动情况较符合,但缺乏实时性且需要较为严格的光照条件。基于运动捕获数据的方法需要价格高昂的设备,且运动灵活性低。Boboc 等提出的基于前向运动学的方法,虽然克服了此问题且具备较强实时性,但角色模型骨骼运动与实际运动之间存在较大偏差。针对以上问题,本文提出了一种新的骨骼动画方法。具体思路如下:首先利用 Kinect 设备捕获的骨骼数据,建模描述人体任意时刻的运动姿态,然后根据人体前后帧运动之间的联系,确定角色模型关节点骨骼旋转矩阵,最后对捕捉对象空间位移变化进行处理,实时调整模型空间位置形成骨骼动画。

按照上述描述,本文提出的骨骼动画方法可概括为以下 4 个阶段:1) 初始姿态假定:对角色模型初始姿态进行固定;2) 单帧姿态建模描述:利用 Kinect SDK 应用程序接口,获取骨骼状态,提取骨骼关节点三维坐标数据,利用骨骼数据对关节点位形进行数学描述,将人体运动姿态变化问题转化为旋转平移问题;3) 骨骼旋转矩阵求解:以关节点单位方向基向量为基础求解骨骼旋转矩阵  $T(j)$ ,利用骨骼旋转矩阵对关节点方向进行调节;4) 根节点位移:更新根节点位置,体现角色模型空间位移变化,得到最终结果。

### 2.2 初始姿态假定

由于不同美工制作的角色模型形态各不相同,因此,模型间必定存在骨骼长短不一致,关节点旋转角度不相等,各个关节点在父节点空间中相对位置不同等差异。这些差异将导致即使对不同角色模型应用相同骨骼数据也无法生成一致的骨骼动画。因此,为了生成符合特定角色模型的骨骼动画,需统一角色模型初始姿态。本文中角色模型的初始姿态设定为直立。

### 2.3 单帧姿态建模描述

Kinect 骨骼数据是非层次、非树形的三维数据,由于各数据之间不存在任何的相关联系,因此,无法直接利用 Kinect 骨骼数据来对角色模型运动姿态进行调整。为了提取蕴含于骨骼数据中运动者的旋转平移信息,本文运用刚体运动学中关于刚体方位的定义来对单帧图形中人体运动姿态进行描述。

为了规定空间中某个关节的方位,建立一直角坐标系  $\{B\}$  与关节点固接。用坐标系  $\{B\}$  的 3 个单位主矢量  $[x \ y \ z]$  相对于参考坐标系  $\{S\}$  的方向余弦组成的三维矩阵来对关节点方位进行描述,记三维矩阵为  $O_S^B$ 。

由迪纳维特-哈坦伯格提出的 D-H 坐标系描述法是目前常用的坐标系建立方法,D-H 坐标系描述法规定:在各个主要部位上固定坐标系,坐标系的  $Z$  轴与关节的轴线重合,而  $X$  轴则是沿着相邻 2 个  $Z$  轴的公垂线, $Y$  轴可由右手坐标系来确定。本文在 D-H 描述法的基础上,结合人体结构特征,提出了一种新的关节点坐标系建立方法,具体步骤如下。

1) 选择在垂直切面上与离待建立坐标系关节点最近的关节点为向量末端,待建立坐标系关节点为首端,建立  $Y$  轴,若垂直切面无最近关节点,则选正  $Y$  轴方向作为关节点  $Y$  轴方向;

2) 在水平切面上选择与待建立坐标系关节点最近的关节点为向量末端,待建立坐标系关节点为首端,建立向量  $AB$ ,若无,则设:  $AB = (1 \ 0 \ 0)$ ;

3) 根据空间解析几何相关知识可知, $X$  轴垂直于  $Y$  轴和  $AB$ ,那么  $X$  轴的方向可由(4)式求得。

$$X = Y \times AB, \quad (4)$$

(4) 式中:  $X = (m_{11} \ m_{21} \ m_{31})$ ;  $Y = (m_{12} \ m_{22} \ m_{32})$ 。

4) 因为  $Z$  轴垂直于以  $X$  和  $Y$  轴为方向的平面,故  $Z$  轴方向可按(5)式求得

$$Z = X \times Y, \quad (5)$$

方向为从  $X$  轴逆时针绕到  $Y$  轴。

按照上述步骤即可得到每个关节点的方位矩阵,设方位矩阵为  $O_B$ ,表示为

$$O_B = [X \ Y \ Z] = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}, \quad (6)$$

(6) 式中,  $Z = (m_{13} \ m_{23} \ m_{33})$ 。

设参考坐标系  $\{S\}$  的 3 个单位向量分别为  $e_1, e_2, e_3$ ,坐标系  $\{B\}$  的 3 个单位向量分别为  $x_1, x_2, x_3$ ,则  $\{S\}$  与  $\{B\}$  之间的方向余弦  $a_{ij}$  为

$$a_{ij} = \cos(\theta_{ij}) = e_i \cdot x_j, \quad (7)$$

由于世界坐标系为标准单位基向量,表示为

$$O_S = [U \ V \ W] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (8)$$

依据(7)式可知,任意关节点的方向矩阵即为坐标系基向量

$$O_S^B = O_B = [X \ Y \ Z] = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}, \quad (9)$$

### 2.4 骨骼旋转矩阵求解

考虑捕捉对象姿态发生变化后的情形,以左臂肘关节向上微微抬起,其他关节相对位置无变化为例,其他运动同理。

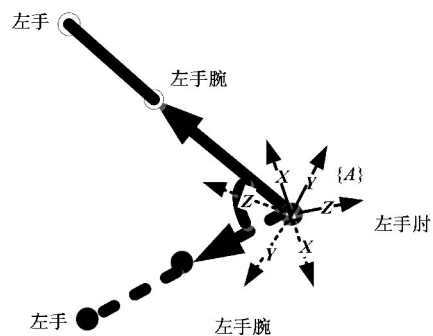


图2 肘关节向上抬起

Fig. 2 Lifting up the left wrist

图2中,实线为后一帧手部位置,虚线为前一帧手部位置。其中,坐标系  $\{B\}$  描述了前一帧左手肘的方位,坐标系  $\{A\}$  描述了后一帧左手肘的方位, $\{B\}$  与  $\{A\}$  有共同的坐标原点,但是方位不同,记  $\{B\}$  为  $P_B = (i_B \ j_B \ k_B)$ ,  $\{A\}$  为  $P_A = (i_A \ j_A \ k_A)$ 。其中  $i_B \ j_B \ k_B \ i_A \ j_A \ k_A$  分别为  $\{B\}$  与  $\{A\}$  的单位主矢量。在关节对应皮肤上取一点,假设为  $P$ ,设  $P$  在两坐标系中的坐标分量分别为  $({}^A P_x \ {}^A P_y \ {}^A P_z)$  与  $({}^B P_x \ {}^B P_y \ {}^B P_z)$ ,根据坐标旋转原理可以得到

$$\begin{pmatrix} {}^A P_x \\ {}^A P_y \\ {}^A P_z \end{pmatrix} = \begin{pmatrix} i_A \cdot i_B & i_A \cdot j_B & i_A \cdot k_B \\ j_A \cdot i_B & j_A \cdot j_B & j_A \cdot k_B \\ k_A \cdot i_B & k_A \cdot j_B & k_A \cdot k_B \end{pmatrix} \begin{pmatrix} {}^B P_x \\ {}^B P_y \\ {}^B P_z \end{pmatrix}, \quad (10)$$

则关节点骨骼旋转矩阵为

$${}^B T = \begin{pmatrix} i_A \cdot i_B & i_A \cdot j_B & i_A \cdot k_B \\ j_A \cdot i_B & j_A \cdot j_B & j_A \cdot k_B \\ k_A \cdot i_B & k_A \cdot j_B & k_A \cdot k_B \end{pmatrix}, \quad (11)$$

则可得关节点骨骼旋转矩阵,结合(2)式可知最终关节点变化矩阵为

$$F(j) = F(p(j)) R(j)^B_A T, \quad (12)$$

## 2.5 根节点调整

由于在运动数据处理阶段忽略了角色模型根节点在世界坐标系下的平移与纵向运动,因此,需要对根节点做额外的位移处理。由于 Kinect 坐标系与游戏引擎坐标系在基本度量单位长度的定义上存在不同,导致直接应用捕捉对象根节点数据到游戏引擎中会产生失真现象,故需对根节点数据进行处理。设 Kinect 与游戏引擎基本度量单位长度差距为  $n$ ,若在 Kinect 世界坐标系中,运动捕捉对象根节点在前后帧之间的位移为  $M_{\text{offset}} = (m_x, m_y, m_z)$ ,则在游戏引擎中角色模型根节点的位移为

$$R_{\text{offset}} = (m_x/n, m_y/n, m_z/n), \quad (13)$$

## 2.6 方法描述

假设角色模型关节点总数为  $n$ 。根据前文描述,本文方法可归纳为以下步骤。

- 1) 检测 Kinect 骨骼状态,若检测到骨骼则进入
- 2) 否则继续检测;
- 2) 置  $i$  为 0,如果  $i < n$  则进入 3),否则转到 6);
- 3) 依据改进的 D-H 坐标系描述法求解  $O_B$ ,依据 (9) 式得到  $O_S^B$ ;
- 4) 使用前后帧方位矩阵,利用 (11) 式求解关节点骨骼旋转矩阵;
- 5)  $i = i + 1$ ,返回 2);
- 6) 利用骨骼旋转矩阵调节关节点;
- 7) 利用 (13) 式调整根节点;
- 8) 结束,重新进入 1)。

从上述描述可知,求解各个关节点的骨骼关节矩阵是本文方法的基本操作。根据系统运行步骤可知,本文方法在一次完整的执行过程中的基本操作执行次数为  $O(n)$ 。

## 3 实验验证

鉴于公开的视频数据库数据量较大,且一般进行了数据压缩存储,存在信息缺失的问题,这会对本文的后续特征点提取产生干扰。因此,本文实验部分以 Kinect 设备检测到的捕获对象骨骼数据作为实验数据来源。为了说明本文方法的有效性,确保对比实验的公平性,共设计了 3 组试验方案,分别为光照对比试验、运动实时性试验和模型运动精度对比试验。

Boboc 等提出的基于前向运动学的方法与本文方法都基于 Kinect 平台且该方法具有较强实时性,同属于国外较新研究成果,故选择该方法作为模型运动精度试验对比对象。光照对比实验主要是在弱光照条件下进行特征点提取对比,基于视觉的方法

特征点提取采用的是目前主流的 surf 算法。

本文硬件实验环境为: Intel Core i3 3.30 GHz CPU,内存 4 GByte,操作系统为 Windows 7,开发环境为 Visual Studio 2010,OGRE 和 Kinect SDK。

### 3.1 强弱光照条件下运动对比实验

图 3、图 4 分别展示了本文方法在正常光照条件下(320 lux)及弱光照条件下(10 lux)的模型运动情况。图 5 给出了二者的关节点识别效果。

浙江大学庄越挺指出:基于视觉的方法在人体运动捕获中,常用的特征点跟踪提取算法对光照条件非常敏感。从图 5 可以看出,鲁棒性较强的特征点跟踪提取算法 surf 在光照强度为 10 lux 时失效,而本文方法在 10 lux 时仍能准确获取人体各关节点的三维坐标,求解骨骼旋转矩阵,实现骨骼动画。从图 5 还可知,本文方法相对于基于视觉的方法,对捕捉对象所穿衣物与背景之间的颜色差别要求更低,色差敏感度较弱,能更佳地应对复杂工作环境。

### 3.2 实时性实验

实时性的实验结果如图 3 所示,每一幅图中右下角方框用于显示人体运动情况,左上方为角色模型实时运动情况。本文方法从提取骨骼数据到角色模型做出相应动作只需 0.007 s,远小于 Kinect 提供的数据时间间隔 0.033 s 和人类视觉停留时间 0.042 s。因此,本文方法能够实时处理 Kinect 提取到的骨骼数据,克服视觉类方法缺乏实时性的问题,且生成的骨骼动画平滑性高,连续性强。

### 3.3 模型运动精度对比实验

Boboc 等以前向运动学为基础,通过前向运动学求解关节点坐标数据,并对坐标数据进行平滑处理,然后使用处理后的数据调整模型,从而产生连贯且平滑的角色运动。图 6 展示了 2 种方法在复杂运动与简单运动 2 种模式下的角色运动效果。其中,简单运动包括双手向上伸直、头部低垂 2 种;复杂运动包括双手叉腰、腿后踢加腰前倾,以及左右手并拢加腰后倾、腿部斜踢 2 种。通过比较可以发现,本文方法中角色模型与捕捉对象的运动更加接近,运动精度更高。

本文以关节点实际位置与计算位置之间的欧拉距离来衡量姿态误差,相应的计算公式为

$$m_d = \max_{i=1}^n (t_s - c_s), \quad (14)$$

$$avg\_error = \frac{\sum_{i=1}^n \frac{(t_s - c_s)}{m_d}}{n}, \quad (15)$$

(15) 式中:  $avg\_error$  表示平均误差;  $n$  代表关节点

总数;  $t_s$  表示关节点真实位置;  $c_s$  表示计算出的位置;  $m_d$  表示最大欧拉距离。利用(14) — (15) 式对每帧结果计算关节点平均误差, 而后对所有帧平均误差取均

值。通过实验测试可知, 本文方法的模型运动误差为 4.2%, 而 Boboc 等提出的方法为 7.4%。因此, 本文方法得到的骨骼动画和实际情况更加相符合。

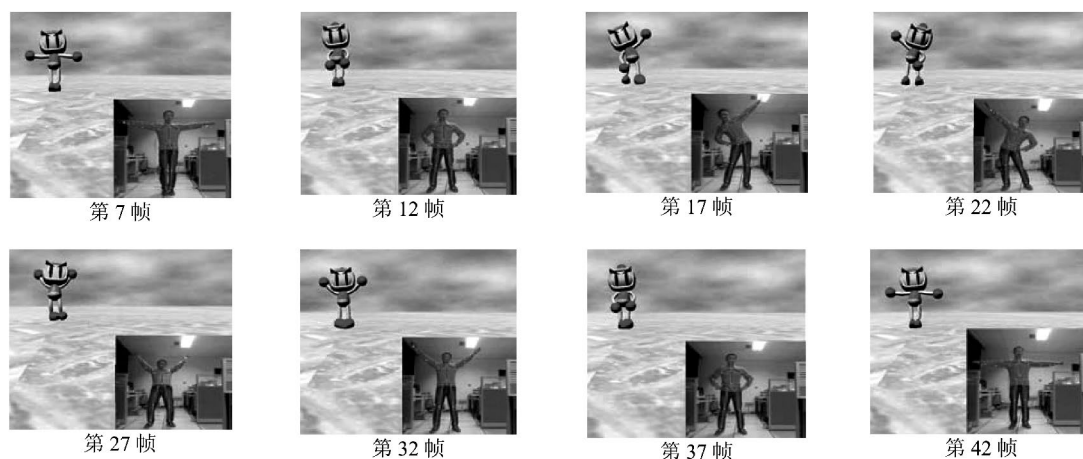


图 3 正常光照条件下实验及实时性效果

Fig. 3 Experimental results under normal light conditions and the real-time effects

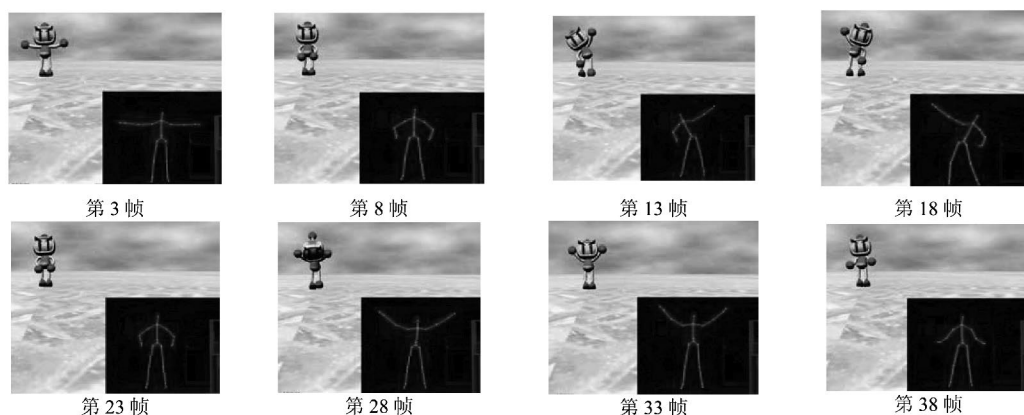


图 4 弱光照条件下实验

Fig. 4 Experimental results under weak light conditions

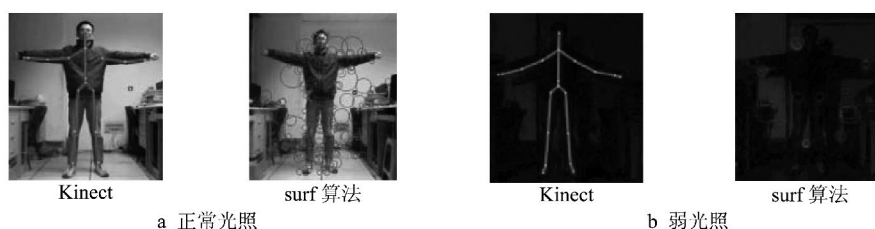


图 5 关节点识别

Fig. 5 Joints identification

## 4 结束语

本文利用 Kinect 设备获取的骨骼数据, 提出了一种新的骨骼动画方法。实验表明, 此方法相比现有解决方法有更强的实时性, 更好的抗光照干扰能

力和更佳的运动精度。目前 Kinect 设备提供的骨骼数据容易出现剧烈抖动现象, 下一步将重点研究如何对 Kinect 设备捕获的数据进行滤波处理, 降低数据抖动带来的误差。

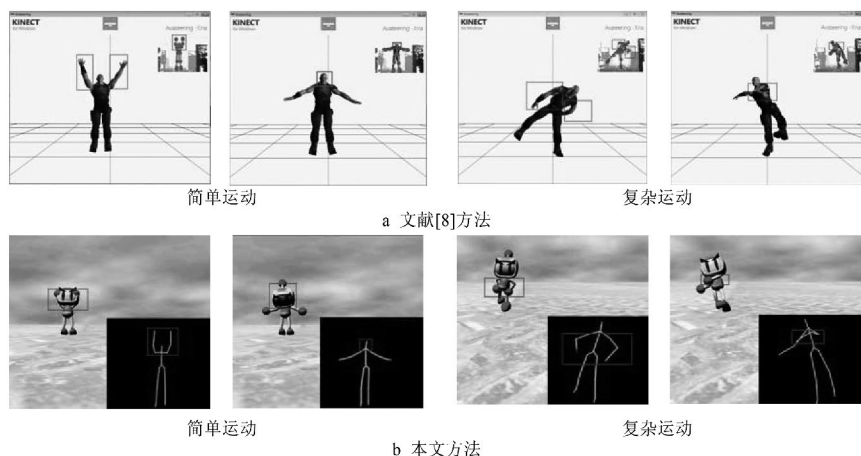


图6 模型运动精度对比

Fig. 6 Comparison of model motion accuracy

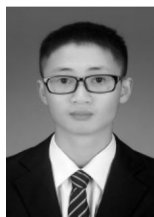
## 参考文献:

- [1] ISMAIL I, SUNAR M S, BASORI A H. Dynamic Motion Control Editor for Virtual Human [J]. International Journal of Interactive Digital Media, 2013, 1(1): 14-17.
- [2] ISMAIL I, KOLIVAND H, SUNAR M S, et al. An Overview on Dynamic 3D Character Motion Techniques in Virtual Environments [J]. Life Science Journal, 2013, 10(3): 846-890.
- [3] 尚华强. 基于 Kinect 的虚拟人物动作仿真研究 [D]. 杭州电子科技大学, 2013.  
SHANG Huaqiang. Simulation study of virtual movement based on Kinect [D]. Hangzhou: Hangzhou Dianzi University, 2013.
- [4] 罗忠祥, 庄越挺, 潘云鹤, 等. 视频中不完全运动特征的跟踪算法 [J]. 计算机辅助设计与图形学学报, 2003, 15(6): 730-735.  
LUO Zhongxiang, ZHUANG Tingyue, PAN Yunhe, et al. Incomplete motion feature tracking algorithm in video sequences [J]. Journal of Computer-aided Design & Computer Graphics, 2003, 15(6): 730-735.
- [5] TONG X, XU P, YAN X. Research on skeleton animation motion data based on Kinect [C] // IEEE. Computational Intelligence and Design (ISCID), 2012 Fifth International Symposium. New York: IEEE Press, 2012, 2: 347-350.
- [6] PARK S I, HODGINS J K. Data-driven modeling of skin and muscle deformation [J]. ACM Transactions on Graphics (TOG), 2008, 27(3): 96.
- [7] CASAS D, TEJERA M, GUILLEMAUT J Y, et al. 4D parametric motion graphs for interactive animation [C] // ACM. Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games. New York, USA: ACM, 2012: 103-110.
- [8] BOBOC R G, TOMA M I, PANFIR A N, et al. Learning new skills by a humanoid robot through imitation [C] // IEEE. Computational Intelligence and Informatics (CIN-TI), 2013 IEEE 14th International Symposium. New York: IEEE Press, 2013: 515-519.
- [9] 况鹰. 基于 Kinect 运动捕获的三维虚拟试衣 [D]. 广东: 华南理工大学, 2012.  
KUANG Ying. 3D virtual try-on based on Kinect motion capture [D]. Guangdong: South China University of Technology, 2012.

## 作者简介:



李红波(1970-),男,甘肃兰州人,高级工程师,硕士生导师,主要研究方向为数字媒体、人机交互、机器视觉。E-mail: lihongbo@cqupt.edu.cn。



冉光勇(1988-),男,重庆人,硕士研究生,主要研究方向为数字媒体、人机交互。E-mail: rangang2007@foxmail.com。



吴渝(1970-),女,重庆人,教授,博士生导师,主要研究方向为网络智能、数字媒体、数据挖掘。E-mail: wuyu@cqupt.edu.cn。



丁林建(1989-),男,安徽安庆人,硕士研究生,主要研究方向为数字媒体、机器视觉。E-mail: 747033572@qq.com。

(编辑: 王敏琦)