

# 双目视觉下人体运动的三维模型描述方法研究\*

魏 玮,董小钊,申媛媛

(河北工业大学 计算机科学与软件学院,天津 300130)

**摘 要:** 提出一种在双目视觉的条件下,利用光流金字塔跟踪算法跟踪关节点,以解决传统光流不能进行大幅度运动跟踪的问题,并利用极线约束进行三维重建获得深度信息,结合对消失关节点的位置预测求出相应的三维坐标,再利用层次化的描述方法将三维坐标转化为三维模型中各个关节点的旋转角度,实现对人体运动的模型描述。实验结果表明,该算法能有效地对人体的运动进行描述。

**关键词:** 人体三维模型;光流金字塔;极线约束;层次化描述

中图分类号: TP301

文献标识码: A

文章编号: 1674-7720(2011)11-0041-04

## Research on 3D model description of human motion in binocular camera system

Wei Wei, Dong Xiaojiao, Shen Yuanyuan

(School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300130, China)

**Abstract:** This paper proposed an approach that in a binocular camera system, to track key points using optical flow pyramid to solve the problem that the traditional optical flow can not solve, and reconstructing the human body to get the depth information by means of epipolar restriction, attached with the prediction of the disappear joints to get the corresponding 3D coordinates, then turned the 3D coordinates to the rotation angles with the means of hierarchical description. Experimental results show the algorithm can track complex human motion and describe.

**Key words:** 3D human model; optical flow pyramid; epipolar restriction; hierarchical description

人体的运动包含了相当丰富的信息,如从人体运动当中可以看出其动作、姿态;根据目前的运动可以推测它在未来某个时刻的位置、速度;可以看出它的表情、神态;可以了解一个人的个性等。正因为如此,对人体运动的研究有着相当广阔的应用空间,如视频监控、运动分析检测、虚拟现实、动画制作等。而其中关于人体关节点的运动跟踪检测以及人体运动姿势的三维重建更是其研究的热门话题。人体姿势的三维重建可以按照视点的个数分为基于单视点的三维重建<sup>[1]</sup>和基于多视点的三维重建。基于单视点三维重建的优点在于它使用方便,可以在大多数的场合中使用,限制的条件较少,只需要一个摄像机或其他输入设备。缺点是由于只有一个视点,得不到相应的深度信息,所以当有自遮挡的情况时会产生匹配的模糊性问题。而基于多视点的三维重建可以部

分解决这个问题。与单视点相比,多视点易于进行三维重建,但多视点的跟踪同样存在特征对应及自遮挡等难点,而且多视点的自动对应也是目前比较有难度的一个课题。其中对于基于多视点的人体姿势三维重建,大体上可以分为利用机器学习的方法和目标跟踪的方法两种。通过机器学习的方法,可以利用大量的先验知识来获得更加准确的有关三维人体姿势的估计。然而,该方法需要大量的样本数据,使用起来十分不方便而且费时,在一定程度上限制了它的应用。目标跟踪的方法主要由两个步骤组成:首先是在每一帧中为每一个特征点定位,并且在每一帧中跟踪相应的特征点;然后通过获得特征点来对人体姿势进行三维重建。本文提出一种在两个摄像机的条件下,利用金字塔光流的方法对关节点进行跟踪检测,并利用极线约束进行三维重建,对运动过程中被遮挡的关节点进行预测处理,最后将求出的

\* 基金项目:河北省自然科学基金(08M006)

关节三维坐标转化成模型中的旋转角,通过层次化的描述,用人体模型来显示相应的动作姿态。

## 1 算法主要思想

### 1.1 三维人体模型

将人体看成是由关节及其连接线段所组成的刚体的集合。如下肢是由膝关节连接的大腿和小腿两个刚体。以一条线段来表示一个刚体,将人体的运动简化为人体框架的运动,这样就得到一个三维人体框架模型<sup>[2]</sup>,如图1、图2所示。该人体模型共包含16个人体关节点,其中所有关节点之间的线段在任何时刻都不会发生形变,而且各个线段之间的长短比例关系可以事先根据个体情况进行适当调整<sup>[3]</sup>。

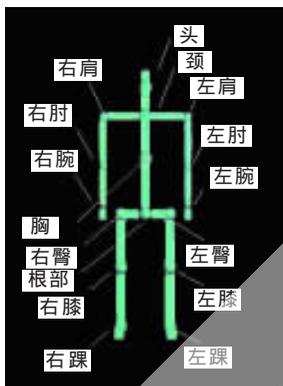


图1 三维人体模型

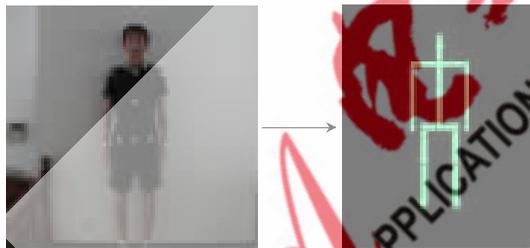


图2 人体关节点对应关系

### 1.2 光流金字塔跟踪算法

传统的 Lucas-Kanade 的光流算法<sup>[4]</sup>主要基于以下三个假设:(1)亮度恒定,即图像场景中目标的像素在帧间运动时外观上保持不变。对于灰度图像,需要假设像素被逐帧跟踪时其亮度不发生变化。(2)时间连续或运动是“小运动”,即图像的运动随时间的变化比较缓慢。在实际应用中指的是时间变化相对图像中目标运动的比例要足够小,这样目标在帧间的运动就比较小。(3)空间一致,即一个场景中同一表面上邻近的点具有相似的运动,它们在图像平面的投影也在邻近区域。但是由于受人体运动和图像噪声的影响,前后帧中对应的特征点的亮点完全相等几乎是不可能的,一方面是人体的运动往往幅度较大,光线难以保证均匀。另一方面由于运动会产生不可避免的遮挡,亮度甚至还会发生突变,因此传统的方法在人体跟踪中受到一定的局限。

本文将对传统的光流方法进行改进,使用一种基于图像金字塔的角点梯度光流计算方法对图像序列中的角点进行亚像素级跟踪<sup>[5]</sup>。该方法的基本思想是:首先构造如图3所示的图像序列的金字塔,金字塔中较高的层是下层图像向下降采样后的图像,原始图像的层数为零。当图像分解到一定的层后,相邻帧的图像间运动幅度就变得足够小,就会满足光流的几个约束条件,此时便可进行光流计算。由高层向低层进行计算,当某一级的光流增量计算出来后,将加到其初始值上,作为其下一层的光流计算初值。这一过程持续进行直到估计出原始图像的光流。

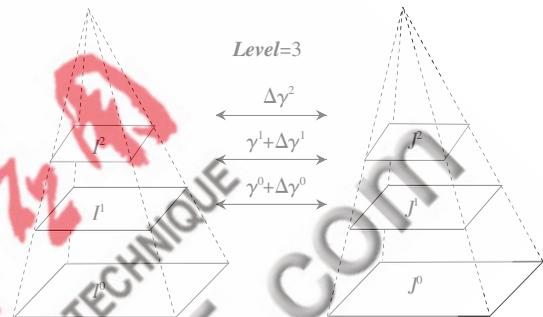


图3 图像序列金字塔

设  $I^0$  为原始图像 ( $w \times h$ ) 的灰度表示,位于金字塔的第0层,并依次生成  $n$  层高斯图像金字塔的第1,2,...,  $n-1$  层,分别表示为  $I^1, I^2 \dots I^{n-1}$ 。对其中的第  $L$  ( $0 < L \leq n-1$ ) 层图像,其灰度计算为:

$$I^L(x, y) = 1/4 I^{L-1}(2x, 2y) + 1/8 [I^{L-1}(2x+1, 2y) + I^{L-1}(2x-1, 2y) + I^{L-1}(2x, 2y+1) + I^{L-1}(2x, 2y-1)] + 1/16 [I^{L-1}(2x+1, 2y+1) + I^{L-1}(2x-1, 2y+1) + I^{L-1}(2x+1, 2y-1) + I^{L-1}(2x-1, 2y-1)] \quad (1)$$

且  $x, y$  仅定义在以下范围:

$$0 \leq x \leq 1/2^L w - 1, 0 \leq y \leq 1/2^L h - 1 \quad (2)$$

对于图像  $I^0$  中的特征点  $x(x, y)$ ,有第  $L$  层中的图像  $I^L$  的  $x^L(x^L, y^L)$  点与之对应,其中  $x^L = x/2^L$ 。由于高层上的误差将向下层传播和放大,因此层数不宜过多。一般地,  $3 \leq n \leq 5$ 。对于给定的两幅连续图像  $I(x, y), J(x, y)$ ,特征跟踪的目标是找到  $I(x, y)$  上一点  $u = (u_x, u_y)$ ,在  $J(x, y)$  上对应具有相似图像强度的另一点  $u+d = (u_x+d_x, u_y+d_y)$ ,其中  $d$  为  $u$  点的平移量。将光流定义为使不同时刻图像区域之间产生最佳拟合的位移,以  $u$  点为中心,在  $I(x, y), J(x, y)$  分别建立  $(2w_x+1, 2w_y+1)$  的对应区域窗口,并假设对应窗口内的像素具有相同的运动位移矢量  $d$ 。定义窗口内像素的灰度方差和跟踪残差  $E(d)$  为:

$$E(d) = E(d_x, d_y) = \sum_{x=u_x-w_x}^{u_x+w_x} \sum_{y=u_y-w_y}^{u_y+w_y} (I(x, y) - J(x+d_x, y+d_y))^2 \quad (3)$$

对图像序列金字塔同一层级  $L$  上的相邻两幅图像  $I^L(x, y), J^L(x, y)$ ,其跟踪残差函数定义为:

$$E^L(d^L) = \sum_{x=u_x-w_x}^{u_x+w_x} \sum_{y=u_y-w_y}^{u_y+w_y} (I^L(x, y) - J^L(x+d_x^L, y+d_y^L))^2 \quad (4)$$

《微型机与应用》2011年第30卷第11期

式中,  $g^L$  为  $L$  层对应的初始光流估计值, 由  $N-1$  层到  $L+1$  层的递推计算得到, 最顶层的初始光流估计值  $g^{N-1}=0$ ;  $d^L$  为  $L$  层计算得到的光流结果, 上下层间的光流计算传递过程为:

$$g^{N-1}=0 \quad (5)$$

$$g^{L-1}=2(g^L+d^L) \quad (6)$$

上述递推过程自上而下进行, 直到第 0 层计算完毕并得到最终的光流计算结果:

$$d=g^0+d^0=\sum_{L=0}^{N-1} 2^L d^L \quad (7)$$

对于运动过程中被遮挡而消失的特征点, 本文采用图像特征点预测方法, 选择不同的位置坐标加权系数, 可以预测出不同的特征点位置。然后根据两个视图中的对应特征点位置重建三维坐标<sup>[6]</sup>, 并利用预测的三维特征点与重建的三维点进行匹配, 从而解决暂时消失的特征的跟踪问题。

由于消失特征随时可能从视域中出现, 因此, 向周围的邻八域搜索是否存在再次出现的人体关节特征, 一旦出现新的特征, 即修改特征的位置和运动属性以指导下帧的跟踪, 并再次利用金字塔光流方法进行跟踪重建。

### 1.3 极线约束

通过立体标定, 可以求出两个摄像机之间的相对位置, 即用  $R, T$  表示旋转位移的关系<sup>[7]</sup>, 由此便可以使用极线约束来描述投影点  $m, m'$  之间的关系。极线约束原理示例图如图 4 所示。在图 4 中,  $o, o'$  与两个平面相交于  $e, e'$ , 三角形  $(o, M, o')$  确定的平面与两个图像平面相交于两条直线  $l(m, e), l'(m', e')$ 。称  $l$  为平面  $I$  上的极线,  $l'$  为平面  $I'$  上的极线。可以看出  $o, o', m, m', M, e, e'$  这些点都在同一个极平面上, 所以  $m$  的对应点  $m'$  一定位于极线  $l'$  上。同理,  $m'$  的对应点  $m$  一定位于极线  $l$  上, 这就是极线约束。利用极线几何约束可大大减少在两幅图像间匹配对应点的计算量。由于  $I$  中点  $m$  在  $I'$  中的对应点在由  $m'$  决定的极线上, 因此搜索空间的维数由二维降为一维, 从而降低了搜索空间。

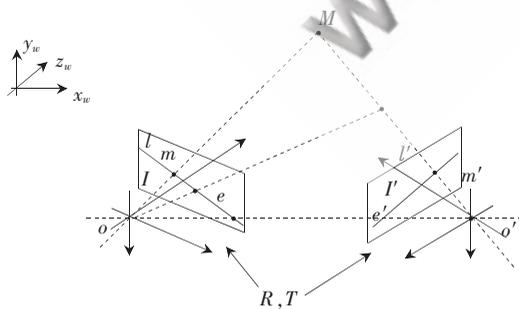


图 4 极线约束原理示例图

### 1.4 层次化描述

对于运动描述, 本文使用一种层次化的运动描述方法<sup>[8]</sup>, 把人体模型看成是类似树的结构, 如图 5 所示。树

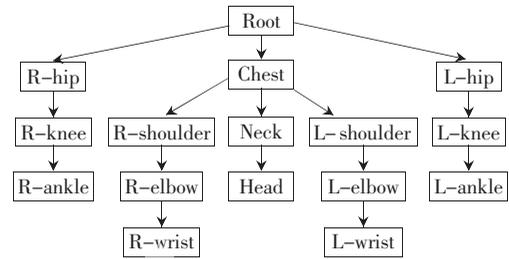


图 5 人体关节的层次关系

的根节点是 Root 点, 其他子节点对应人体模型中的各个关节点。整个人体的运动可以看成是由平移和旋转组成的, Root 点的旋转决定人体模型所对应的方向, 人体的平移即为 Root 点的位移量, 其他各节点的旋转是在以父节点为坐标原点的相对坐标系下的旋转。各个关节点的位置可以根据各个关节点间的线段长度和旋转向量求出, 而关节点间的线段长度可作为先验知识获得。图中的节点 R\_ankle, 其位置与旋转平移向量的对应关系为:

$$\vec{P}_{R\_ankle}(x, y, z) = \vec{T}_{Root} \vec{R}_{Root} \vec{T}_{R\_hip} \vec{R}_{R\_hip} \vec{T}_{R\_knee} \vec{R}_{R\_knee} \vec{T}_{R\_ankle} \vec{P}_0(x, y, z) \quad (8)$$

其中,  $\vec{P}_{R\_ankle}(x, y, z)$  表示 R\_ankle 点的世界坐标;  $\vec{P}_0(x, y, z)$  表示初始位置时关节点 R\_ankle 在以父节点 R\_knee 为原点的相对坐标系下的相对坐标;  $\vec{T}_i (i=Root, R\_hip, R\_knee, R\_ankle)$  表示节点  $i$  绕父节点的旋转向量, 旋转向量由绕  $x, y, z$  轴的三个旋转分量组成, 其中两个分量决定关节点的位置, 另外一个为扭矩分量, 决定该段骨架的方向。通过三维重建可以求出视频中人体关节点的三维坐标, 可以先计算关节树的根节点的平移和旋转向量, 然后逐步求各个子节点的旋转向量。根据式(8)的关系可以求出视频序列每一帧人体模型各个关节点对应的旋转向量<sup>[9]</sup>。

根据上述方法可求得各关节点的旋转和节点的平移向量。将视频序列中人体运动描述为:

$$M(t) = (T(t), R_1(t), \dots, R_n(t)) \quad (9)$$

其中,  $T(t)$  表示  $t$  时刻点 Root 的平移向量,  $R_i(t)$  表示  $t$  时刻第  $i$  个关节点的旋转向量。

在已经知道每个关节点的运动除了包括绕父节点旋转外, 还包括随父节点绕祖父节点旋转和平移。因此, 三维数据向旋转平移数据转化的关键是在各个非叶子节点处建立以该非叶子节点为原点的参考坐标系。以人体上臂为例来说明旋转数据的计算方法。由于三维数据中没有考虑扭矩的作用, 可假设绕轴的运动分量为 0; 假设旋转方向为先绕  $z$  轴旋转  $\beta$ , 再绕  $y$  轴旋转  $\theta$ 。然后, 将三维数据  $(x, y, z)$  变换到上臂参考坐标系下, 其坐标为  $(m_1, m_2, m_3)$ , 则参考坐标系下的三维数据  $(m_1, m_2, m_3)$  与参考坐标系下的上臂初始位置  $(-l, 0, 0)$  之间的关系可以描述为<sup>[10]</sup>:

$$\begin{bmatrix} \cos\beta & -\sin\beta & 0 & 0 \\ \sin\beta & \cos\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -l \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ 1 \end{bmatrix} \quad (10)$$

根据式(10),可得到以下方程组:

$$\begin{cases} -l \cos\theta \cos\beta = m_1 \\ -l \cos\theta \sin\beta = m_2 \\ l \sin\theta = m_3 \end{cases} \quad (11)$$

根据式(11)方程组,可以求出参考坐标系下,上臂绕z轴和绕y轴的两个旋转角分别为:

$$\begin{cases} \beta = \arctan m_2 / m_1 \\ \theta = \arcsin m_3 / l \end{cases} \quad (12)$$

对于右下臂,假设先绕y轴旋转 $\beta$ ,然后绕z轴旋转 $\theta$ ,则参考坐标系下三维数据 $(n_x, n_y, n_z)$ 与初始位置 $(-l, 0, 0)$ 之间的关系可以描述为:

$$\begin{bmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -l \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} n_x \\ n_y \\ n_z \\ 1 \end{bmatrix} \quad (13)$$

根据式(13),可以求出两个旋转角分别为:

$$\begin{cases} \beta = -\arctan n_z / n_x \\ \theta = -\arcsin n_y / l \end{cases} \quad (14)$$

以此类推,可以计算出人体各个关节的旋转分量,其实现效果如图6所示。

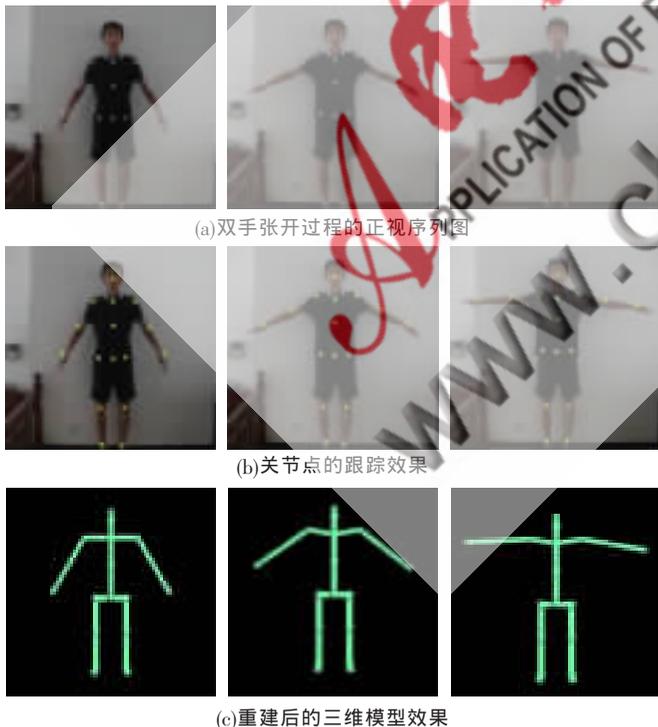
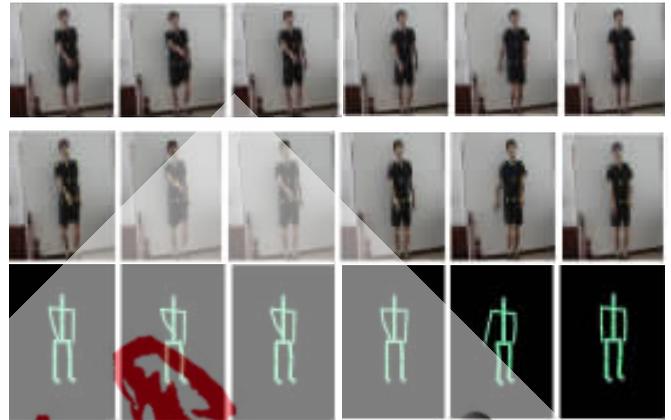


图6 人体各个关节的旋转分量实现效果

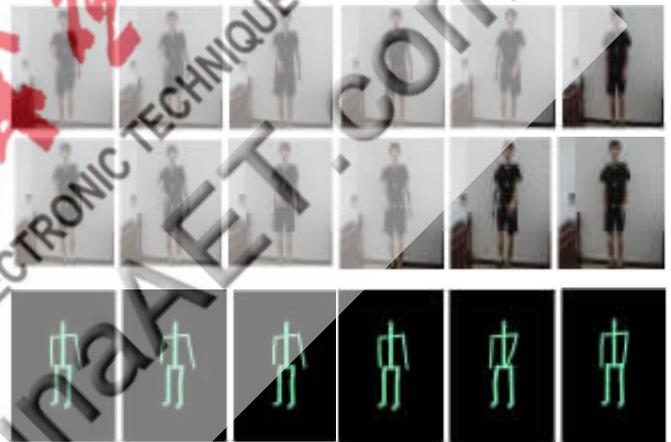
## 2 实验结果

根据本文提出的算法,使用VC++开发一个人体运

动跟踪描述系统,可运行在Windows平台上。实验中的视频是使用两个具有平行光轴的摄像机拍摄而成。图7是拍摄的一段图像序列的跟踪结果以及恢复出的三维骨架序列。



(a)左侧摄像机拍摄的行走姿态及跟踪重建效果



(b)右侧摄像机拍摄的行走姿态及跟踪重建效果

图7 拍摄的一段图像序列的跟踪结果

本文提出一种双目视觉下的三维人体运动描述算法。该算法鲁棒性较高,因为除利用贴在人体关节处的用于构建三维标记点的白色标记点之外,没有利用图像上的其他特征。实验结果表明,该算法能有效地对人体运动进行描述,并且当有比较大幅度的运动及遮挡的情况时,依然可以较好地跟踪和预测关节点位置,并进行重建显示。未来的工作是要加入更多的人体约束和运动学约束以消除动作的多义性,进一步提高描述的准确性和鲁棒性。

## 参考文献

- [1] Zou Beiji, Chen Shu, Cao Shi, et al. Automatic reconstruction of 3D human motion pose from uncalibrated monocular video sequences based on markerless human motion tracking[J]. Pattern Recognition, 2009,42:2.
- [2] 刘小明,庄越挺,潘云鹤.基于模型的人体运动跟踪[J].计算机研究与发展,1999,36:10.
- [3] SHEN B C, SHIH H C, HUANG C L. Real-time human

- motion capturing system[C]. Proceedings of IEEE International Conference on Image Processing, 2005:11-14.
- [4] 张永亮,卢焕章,高劼,等.一种改进的 Lucas-Kanade 光流估计方法 [J]. 海军航空工程学院学报,2009,24(4): 444-445.
- [5] 陈乐,吕文阁,丁少华.角点检测技术研究进展[J].自动化技术与应用,2005,24(5):1-4.
- [6] 张娟,潘建寿,吴亚鹏,等.基于双目视觉的运动目标跟踪与测量 [J]. 计算机工程与应用,2009,45 (25):191-193.
- [7] 蔡涛,李德华,朱洲,等.基于彩色图像序列的特征检测和跟踪[J].计算机工程,2005,31(8):12-13.
- [8] Chen Yisheng, LEE J, PARENT R, et al. Markerless monocular motion capture using image features and physical constraints [C]. Proceedings of Computer Graphics International. Washington DC: IEEE Computer Society, 2005:36-43.
- [9] KIM S M, PARK C B, LEE S W. Tracking 3D human body using particle filter in moving monocular camera[C]. Proceedings of the 18th International Conference on Pattern Recognition. 2006:805-808.
- [10] SILAGHI M, PLAENKERS R, BOULIC R, et al. Local and global skeleton fitting techniques for optical motion capture [C]. Proceedings of International Workshop on Modelling and Motion Capture Techniques for Virtual Environments. London: Spinger-Verlag,1998:26-40.

(收稿日期:2011-01-04)

## 作者简介:

魏玮,男,1960年生,教授,主要研究方向:模式识别,机器视觉,计算智能。

董小铨,男,1984年生,硕士研究生,主要研究方向:三维重建,视频图像处理。

申媛媛,女,1986年生,硕士研究生,主要研究方向:计算机视觉,图像处理。