

无需人工干预的关节点提取和三维重建

魏 玮 王丹丹 刘 静 刘 命

(河北工业大学计算机科学与软件学院 天津 300401)

摘 要 随着现在人体的运动捕获和行为理解的研究的发展,对这项研究有了越来越高的要求。相对于原来的手动提取人体关节点作为特征点来研究,如何使得提取特征点更加自动化,对以后的运动捕获和行为理解的研究意义重大。提出一种在单目视觉条件下在第一帧自动提取人体关节点位置的方法,来解决传统的以手动标定提取人体关节点的问题,并且利用光流稀疏 L_K 算法对提取出的关节点进行运动跟踪,得到运动人体二维坐标信息,结合像机模型通过几何计算获得人体关节点的深度信息。

关键词 关节点,比例正交投影,比例因子,重建

中图分类号 TP391 **文献标识码** A

Articulation Points Extraction without Manual Intervention and 3D Reconstruction

WEI Wei WANG Dan-dan LIU Jing LIU Ming

(School of Computer Science and Software, Hebei University of Technology, Tianjin 300401, China)

Abstract Now with the development of the research of the motion capture and the behavior understanding of the human, the demand is increasing. The people extracted the articulation points as the feature points manually before, compared with that. How to make the extraction of the feature points much more automated is meaningful to the research of the motion capture and the behavior understanding later. This paper proposed the method of extracting the dynamic human's articulation points and obtaining the position of the articulation points on the first frame image automatically from monocular video sequences to solve the disadvantage of traditional method about extracting the points from marked human body manually. Then the points were tracked with the light flow sparse L_K algorithm to acquire the information of two-dimensional coordinate about the moving human. Finally, combining the camera model the relative depth of the points can be obtained through geometric calculation.

Keywords Articulation points, Scaled-orthographic camera model, Scale factor, Reconstruction

1 引言

视频人体运动分析是计算机研究领域的重要研究方向,主要内容是从单个或多个像机拍摄的视频序列中获取人体运动信息,进而进行三维重建。三维重建的研究已经比较成熟,其中文献[1]提出的三维重建是利用 Harris 或者 DoG 算法提取特征点再进行三维重建的,最早人们是在人体的关节位置放置反光的硬件设施,而罗忠祥^[2]在对于运动人体重建的研究中通常是将特殊的 marker 贴在运动者的关节上,然后利用特殊的硬件生成一组运动数据。还有庄越挺等人^[3,4]是对第一帧手动提取关节点的位置,有人工的干预。而本文提出的对人体的运动分析不需要利用 Harris 或者 DoG 算法提取角点,对于在关节点上贴 marker 的方法有所改进,不需要任何标记并且在第一帧也不需要人工干预自动地提取人体的关节点,在第一帧图像利用本文提出的算法自动地计算出了关节点的位置坐标。目前,基于多摄像机有标记点的三维人体姿

态捕获技术已经较为成熟并在电影特技、游戏和动画制作等领域获得了广泛的应用,但是由于是基于多摄像机,对设备的要求高,并且需要摄像机标定,像机标定也增加了算法的成本。并且是手动标记关节点,缺乏研究的自动性,还有就是人们提取人体的立体特征是为了研究人体的运动,而研究人体运动只需要知道人体关节点之间的相对信息。所以为了解决以上问题,本文提出了单目的关节点自动提取的人体立体特征的获取方法。系统流程图如图 1 所示。

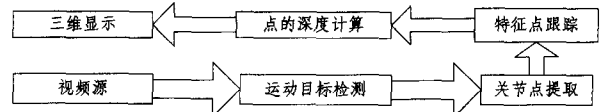


图 1 系统流程

2 关节点自动提取

为了实现自动提取人体关节点,需要将运动人体的前景

到稿日期:2012-06-01 返修日期:2012-09-05

魏 玮(1960-),男,教授,主要研究方向为模式识别、机器视觉、计算智能;王丹丹(1986-),女,硕士生,主要研究方向为计算机视觉、图像处理, E-mail: computer-wang@163.com;刘 静(1985-),女,硕士生,主要研究方向为计算机视觉、图像处理;刘 命(1987-),女,硕士生,主要研究方向为计算机视觉、汉字手写体识别。

从图像中提取出来,再利用本文提出的增加人体身高的约束来自动标定人体的关节点作为特征点。本文自动标定特征点包括目标分割和关节点自动标注。本文的人体模型用有 16 个关节点的人体模型表示,如图 2 所示。

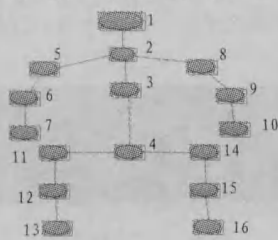


图 2 人体骨架模型

2.1 运动人体目标图像分割

本文采用背景差分法来分割运动人体,背景差分是目前运动分割中最为常用的一种方法,一般能够提供比较完整的特征数据。

2.2 关节点自动提取

人体运动的研究领域中,由于在第一帧手动标注人体关节点^[3,4],因此在人体的运动分析当中可以准确地标示出人体的二维关节点的坐标,此方法简单可行。还有人事先在人体的关节点处贴上标示物或者高亮物体,然后在标签处手动标注,使得关节点提取的自动化受到了限制。福州大学计算机图像图形研究所的陈国栋等人^[5]利用了人体特点和黄金比例来进行骨架提取。本文提出了增加医学生理上人体身高的约束^[6](见图 3)来自动提取人体关节点,在第一帧图像利用本文提出的算法自动地计算出了各个关节点的位置坐标,完全不需要人工干预。具体方法如下所述。

运动目标检测后,获得人体区域的轮廓点坐标,以 Freeman 链码^[7]的方式表示,将人体区域划分为若干 Apar 区域,并且对人体轮廓矢量化,拟合直线^[8]。根据人体轴对称的特点,要求用于提取关节点的第一帧基本满足身体轴对称的特点,算法的流程如下:

输入:人体的轮廓点坐标(即检测出的运动目标)

中间过程:

- 从上到下扫描分割出来的前景图片,并记录纵坐标的最大值 y_{max} 和最小值 y_{min} 。两者之差即为图 3 所示的 H (这样根据图 3 可以将身体各部分的长度大概计算出来)。
- 从左到右扫描分割出的前景图片,记录横坐标的最大值 x_{max} 和最小值 x_{min} ,两者和的平均值即为人体中心横坐标。
- 颈部、肩关节和髋关节等这样的关节点的坐标只需要利用其相对应的人体比例,例如寻找左肩关节的坐标值,根据步骤 b)求出了人体中心的横坐标,用这个数减去 $0.1259H$ 就是左肩关节的横坐标,再用纵坐标的最小值加上 $0.182H$ 所得即为肩关节的纵坐标,用同样方法计算出其他点的坐标。
- 寻找头部、胸部、肘、腕、膝和踝关节的坐标时,比如找膝关节 p 的坐标,则 $y_p = y_{min} + 0.285H$,再找 x_p ,过 y_p 与人体矢量化后的轮廓有 4 个交点,将人体轮廓分为左半轮廓和右半轮廓两部分,在找左膝关节点时候,遍历与左半轮廓交点的连线上面的所有点,计算每个点到矢量化的轮廓直线的距离,分别记作 l_1 和 l_2 ,计算 l_1 和 l_2 的差值,如果小于给定阈值,则该点坐标即为所求 p 点坐标。

输出:各个关节点的坐标值。

结束。

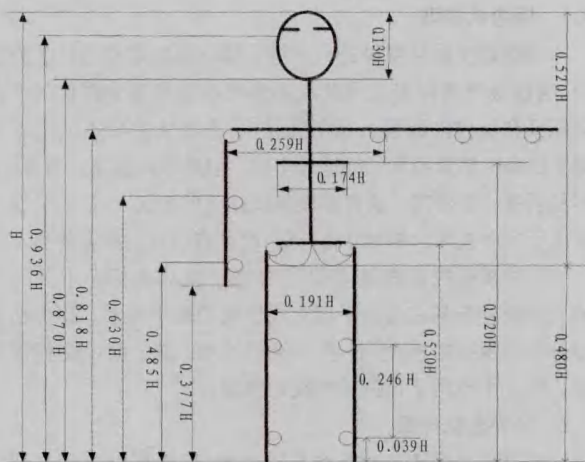


图 3 人体各部分比例图

结果如图 4 所示。

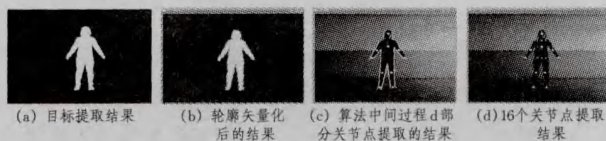


图 4 关节点提取结果图

3 三维坐标计算

实验数据涉及 699 帧自己拍的图片,在提取出视频第一帧的关节点之后,利用光流稀疏 L_K 算法^[9]跟踪人体的 14 个关节点,这样对于人体运动过程中的每一帧的关节点的二维坐标都可以得到,跟踪结果如图 5、图 6 所示。



图 5 自动标注关节点跟踪效果图



图 6 手动标注关节点跟踪效果图

图 5 表明,利用本文算法提取的关节点跟踪效果比较好,为客观验证本文算法的精确度,计算了手动标定点与本文算法得出的关节点间的距离,将其定义为关节点位置误差。计为:

$$e_{fi} = |J_{fi}(x, y) - J'_{fi}(x, y)| \quad (1)$$

式中, e_{fi} 是第 f 帧第 i 个关节点的位置误差; $J_{fi}(x, y)$ 是本文算法得到的关节点坐标, J'_{fi} 是手动标定得到的关节点坐标。 f 取值为 $1 \leq f \leq 460$, 每隔 50 帧取一帧。 $1 \leq i \leq 16$, 引入关节点位置误差率为 $\eta_m, \eta_n = e_f/s$, 其中 $s = \max(M, N)$, M 和 N 分别表示单帧图像的行和列。取其中 10 帧关节点最大误差率如表 1 所列。

表 1 关节点位置最大误差率(%)

关节点序号	1	2	3	5	6
最大误差率	4.3	4.57	3.51	2.74	2.11
关节点序号	7	10	12	14	15
最大误差率	3.06	3.8	2.66	3.50	2.07

3.1 摄像机模型

目前基于单目视频进行三维人体特征提取和估计的方法中,通过摄像机标定技术获取摄像机参数来建立图像的平面坐标系和世界坐标系关系的方法^[10]易于从关节点的二维图像平面坐标估算出其三维世界坐标。但是像机标定计算复杂,不易于进一步研究。这样非摄像机标定的方法^[11]目前只是计算出了运动人体的相对立体特征,因此有很好的研究价值。

摄像机的投影成像模型^[12]描述了世界坐标系下某一点的三维坐标与其在图像平面上投影间的数学关系。摄像机模型分为透视投影模型(针孔摄像机模型)和比例正交投影模型。本文中使用了比例正交投影模型。

3.2 三维坐标计算

当目标和摄像机的距离足够大时,目标反射出的光线几乎是平行光。当焦距 $f \rightarrow \infty$,使得 $f/z_c \rightarrow 1$ 。根据像机的成像模型有公式为:

$$\begin{cases} u = f \frac{x_c}{z_c} \\ v = f \frac{y_c}{z_c} \end{cases} \quad (2)$$

根据式(2)得到 $u = x_c, v = y_c$ 。当目标位置接近于摄像机光轴且目标表面各点与摄像机之间的平均距离 \bar{z} 远大于目标自身维度时也就满足了正交投影模型的条件,即 $\bar{z} \gg dz$,根据式(2)可以得到目标任意关节点的空间三维坐标 (x, y, z) 与其投影点在图像坐标系下的二维坐标 (u, v) 间的关系为: $u = fx/z \approx f\bar{x}/\bar{z}, v = fy/z \approx f\bar{y}/\bar{z}$ 。令 $s = f/\bar{z}$ 表示比例因子,可以表示为:

$$\begin{bmatrix} u \\ v \end{bmatrix} = s \begin{bmatrix} 100 \\ 010 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3)$$

经过运动人体跟踪之后,得到了 14 个关节点的二维坐标值 $\bar{X}_k = (u_k, v_k), k=1, \dots, 14$,将其作为输入,根据比例正交投影模型求取各个关节点对应的相对三维坐标值 $\bar{X}_k = (x_k, y_k, z_k)$,根据上面的式子推导可以得:

$$\begin{cases} x_k = \frac{u_k}{s} \\ y_k = \frac{v_k}{s} \end{cases} \quad (4)$$

式中, u_k 和 v_k 是已知量。因此只需知道 s 便可得到 x_k 和 y_k ,而深度信息 z_k 需要肢体长度信息。令 l_{mn} 表示关节点 m 的肢体段长度,由空间点距离公式得:

$$l_{mn}^2 = (x_m - x_n)^2 + (y_m - y_n)^2 + (z_m - z_n)^2 \quad (5)$$

由式(4)、式(5)得到关节点 m 和 n 的相对深度为:

$$\Delta z = |z_m - z_n| = \sqrt{\frac{l_{mn}^2 - ((u_m - u_n)^2 + (v_m - v_n)^2)}{s^2}} \quad (6)$$

由式(6)来看,有了肢体长度,两关节点的相对深度是关于比例因子 s 的函数。令图像中两关节点之间的以像素为单位的初始化肢体长度为 d_{mn} ,由比例正交投影关系得:

$$l_{mn} = \frac{d_{mn}}{s} \quad (7)$$

根据式(6)、式(7)得:

$$\Delta z = |z_m - z_n| = \frac{\sqrt{l_{mn}^2 - ((u_m - u_n)^2 + (v_m - v_n)^2)}}{s} \quad (8)$$

由式(8)可以看出只要确定了比例因子 s ,以关节点 m 或

者 n 中任意一点作为参考点,便可以求出另一点的相对三维坐标值,并且在对三维人体运动行为进行分析时,只要知道关节点之间的相对坐标值即可。这里把人体的模型表示成了一个树模型,选择了腹部关节点 $\bar{X}_{root} = (x_{root}, y_{root}, z_{root})$ 作为根节点,并将其放置于参考平面上 ($z_{root} = 0$)。这样利用 opengl 将三维人体显示出来。显示结果如图 7 所示。



图 7 三维人体显示

结束语 本文提出了一种结合医学生理解剖学中人体身高比例在首帧自动提取人体关节点的方法。该方法不需要人工干预,可以用于对关节点的跟踪进而进行运动人体的三维重建以及运动捕获和行为理解,使得提取关节点不再局限于必须在人体运动的首帧标定关节点的位置或者是在运动者的关节点上贴标签,而是更加自动化,其要求满足第一帧人体完全轴对称。也可以用于对提取出的关节点进行跟踪,以获取立体特征。实验结果表明,本文提出的方法使得运动人体的研究更加趋于自动化,对于运动人体的行为理解的研究有一定的价值。由于 L_K 跟踪算法受光照影响特别严重,未来的工作是对自动提取的关节点的跟踪进一步的研究,使得此算法更加完善。

参考文献

- [1] Furukawa Y, Ponce J. Accurate, dense and robust multi-view stereopsis[J]. IEEE Pattern Analysis and Machine Intelligence, 2008, 1(1): 1-14
- [2] 罗忠祥. 视频流中的人体运动提取与运动合成[D]. 杭州: 浙江大学, 2002
- [3] 庄越挺, 刘小明, 潘云鹤, 等. 运动图像序列的人体三维运动骨架重建[J]. 计算机辅助设计与图形学学报, 2000, 4(12): 245-250
- [4] 雷涛, 罗薇薇, 樊养余, 等. 复杂环境下的运动人体骨架提取算法[J]. 计算机应用研究, 2010, 8(27): 3194-3197
- [5] 陈国栋, 李建微, 潘林, 等. 基于人体特征三维人体模型的骨架提取算法[J]. 计算机科学, 2009, 7(36): 295-297
- [6] Yoo J-H, Nixon M S, Harris C J. Extracting Human Gait Signatures by Body Segment Properties [C]// Proceedings of IEEE Southwest Symposium on Image Analysis and Interpretation. Southampton, 2002
- [7] 于国防, 王莉. 压缩型顶点链码的研究[J]. 中国图像图形学报, 2010, 15(10): 1465-1470
- [8] 宫宇. 运动人体二维特征点间的与三维重建[D]. 大连: 大连理工大学, 2009
- [9] 张永亮, 卢焕章, 高劫, 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
- [10] Zhao J, Li L. Human Motion Reconstruction from Monocular Images using Genetic Algorithms[J]. Computer Animation and Virtual World, 2004, 15(1): 407-414
- [11] Bregler C, Malik J. Tracking people with twists and exponential maps[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Santa Barbara, CA, 1998: 8-15
- [12] 马颂德. 计算机视觉——计算理论与算法基础[M]. 北京: 科学出版社, 1997