

中图分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2010)12-1804-07

论文索引信息: 徐萧萧, 王智灵, 陈宗海. 视频序列中基于头肩分割的人体位姿估计算法[J]. 中国图象图形学报, 2010, 15(12): 1804-1810

视频序列中基于头肩分割的人体位姿估计算法

徐萧萧, 王智灵, 陈宗海

(中国科学技术大学自动化系, 合肥 230027)

摘要: 当前常用3维重构的方法表示和计算视频中的人体位姿,但由于这些方法通常需要多个摄像头,不仅限制条件多,且计算复杂度高,为此,提出了一种基于头肩分割的人体位姿估计算法。该算法首先对视频中的人体进行头肩定位,然后利用人体头部的平面成像特点计算头部位姿,同时利用人体肩部的轮廓变化特点计算躯干位姿;最后结合头部和躯干的位姿估计运动中的人体位姿。实验结果证明,该算法是有效和优越的。

关键词: 人体位姿估计; 行为识别; 计算机视觉; 视频序列

Human pose estimation algorithm based on segmentations of human head and shoulder in video sequences

XU Xiaoxiao, WANG Zhiling, CHEN Zonghai

(Department of Automation, University of Science and Technology of China, Hefei 230027)

Abstract: To describe and estimate a human pose in video sequences, most methods use 3D reconstruction. These methods often need multiple cameras, have high computation complexity and need many limit conditions. Because of these limitations, this paper proposes a human pose estimation algorithm based on segmentations of human head and shoulder. In this algorithm, the positions of human head and shoulder are located, then the head pose is obtained from the characteristics of plane imaging of human head and the body pose is obtained from the characteristics of the silhouette of human shoulder. Moving human pose is estimated by both head and body poses. The experiment results demonstrate the validity and superiority of the proposed algorithm.

Keywords: estimation of human pose; motion recognition; computer vision; video sequences

0 引言

行为识别是一种对人的运动模式进行分析,并用自然语言加以描述的视觉处理技术。近年来由于其视频监控、人机交互、视频会议、基于内容的图像存储和检索以及虚拟现实等多方面有着广阔的应用前景,使得越来越多的研究人员开始关注这一领域的研究^[1],例如在机场、银行等场合,安保人员希望计算机视觉监控系统能够自主的检测出人们潜在

的危险行为,以排查可疑人员,杜绝安全隐患;又如在体育运动中,通过有效的行为识别算法分析运动员的动作,能够帮助选手提高训练效率。人的行为识别研究越来越受到重视,主要的方法有基于模板的方法^[2]、基于概率网络的方法^[3-5]和基于文法技术的方法^[6]。其中,基于概率网络的方法应用最为广泛,其特点是先将人的每一种静态姿势定义成一个状态,并将人的行为定义为这些状态的遍历过程,状态与状态间的转换用概率的方法表示,然后通过计算联合概率密度来实现行为的识别。可见合理的

基金项目: 国家高技术研究发展计划(863)基金项目(2007AA04Z227)。

收稿日期: 2009-03-03; 改回日期: 2009-06-16

第一作者简介: 徐萧萧(1984—),男,中国科学技术大学自动化系博士研究生。E-mail: telly@mail.ustc.edu.cn。

姿态表示和计算方法是行为识别的基础和必不可少的环节。由于人体是一个多结构、姿态可变且存在大量自遮挡的目标,因此其姿态表示和计算面临许多困难。人体姿态可分为动作姿态和体位姿态(以下简称“位姿”),动作姿态是指人在某一个时刻的动作行为,位姿是指人体在 3 维空间的朝向。当前,人体姿态的表示和计算方法主要可分为基于表观的方法、基于人体模型的方法和基于 3 维重构的方法 3 类:其中,基于表观的方法不对人的物理结构进行直接建模,而是采用目标的颜色信息、轮廓信息以及外观形状等对人体姿态进行分析,如 Yamato 等人采用 2 维网格特征来表示人体姿态^[7],即首先在图像中检测出人体区域,然后将整幅图像分割成若干网格,再计算每个网格中人体像素数占整个网格像素数的比例,并以此表示人体姿态,这类方法通常用于分析人的动作姿态,由于其仅仅利用了 2 维图像中的表观信息,所以难以估计人体位姿。基于人体模型的方法^[8]是先使用线图模型、2 维模型或 3 维模型来对人体进行建模,然后通过分析这些参数化的人体模型来估计人体姿态,这类方法通常用来分析人的肢体位置以及表示一些比较精细的动作,其对图像分辨率和目标检测的精度要求较高。基于 3 维重构的方法^[9]是先将 2 维运动目标重构成 3 维运动目标,通常需要多摄像头在不同位置拍摄目标图像,并在这些图像中找出相对应的特征点;然后利用摄像头参数和成像公式计算出目标在 3 维空间中的相对坐标,这类方法从理论上说,能够估计人体位姿,但其计算复杂度较高,且假设限制条件较多。文献[10]给出了一种基于 3 维重构的人体位姿估计算法,但它必须使用人的背部轮廓图像,并且实验时还需要人体双臂水平张开。

基于表观的方法和基于人体模型的方法一般用来分析动作姿态,而基于 3 维重构的方法虽然可以得到人体位姿,但由于其需要事先标定摄像头参数,且计算复杂度较高,故实际应用时受到很大限制,于是本文提出了一种基于头肩分割的位姿估计算法。由于人体头部相对躯干可以自由转动,头部位姿与躯干位姿可能出现不一致的情况,所以与传统方法不同,本文将人体位姿分为头部位姿和躯干位姿两部分,并分别进行计算。由于人体头部区域的颜色分布和肩部的轮廓特征随着人体姿态的变化呈现出一定的变化规律,因此可通过分析人体头部区域的颜色分布特点和肩部的轮廓特征来分别计算头部位

姿和躯干位姿。在此基础上,再分析人在运动时头部位姿和躯干位姿的关联性,并定义了一套位姿判别规则,最终实现了视频中的人体位姿估计。与前人方法相比,本文算法主要有以下两个特点:1) 仅需单摄像头,且无需预先标定摄像头参数,就能实现人体位姿估计;2) 实时性较好。

1 问题描述和系统框架

为了计算人体位姿,首先要对其进行定义,这里将人体位姿定义为人体的朝向与摄像头光轴方向的夹角。该角度理论上的范围为 $0 \sim 360^\circ$,通过前面的分析可知,要精确计算这个角度是十分困难的,因而可将问题进行一定的简化。本文将人体位姿定义为一个包含 6 个状态的有限状态空间,即 $S = (s_1, s_2, s_3, s_4, s_5, s_6)$,其中 s_1 表示人面向摄像头, s_2 表示人背向摄像头, s_3 表示人面向摄像头向右侧身, s_4 表示人面向摄像头向左侧身, s_5 表示人背向摄像头向右侧身, s_6 表示人背向摄像头向左侧身,于是问题就简化为识别人体位姿在该状态空间中所对应的状态问题。该方法的基本框架如图 1 所示。

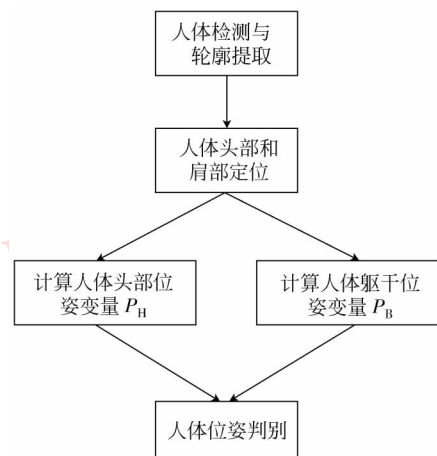


图 1 系统框架

Fig. 1 Architecture of system

在计算人体位姿时,通常需要对人体姿态进行一定的先验假设,例如文献[10]中假设人体双臂水平张开。本文假定人体为直立行走的状态,摄像机拍摄的俯仰角度不超过 30° 。为计算人体位姿,需首先检测出图像中的人体区域,并提取其轮廓。文献[11]给出了进行人体区域检测、轮廓提取以及归一化操作的具体方法,这些方法具备良好的鲁棒性,

在不同的背景环境下,能够准确地得到人体区域和轮廓。本文采用文献[11]中的方法来进行人体检测,在获得归一化的人体轮廓后,先对其进行数学建模,同时利用轮廓特征实现人体头肩的定位;然后分别计算人的头部位姿和躯干位姿。这里定义了两个变量 P_H 和 P_B , P_H 表示头部位姿, P_B 表示躯干位姿;最后通过位姿判别规则得到人体位姿。

2 人体位姿估计算法

2.1 人体轮廓特征的数学建模

本文通过 3 个拟合函数对人体轮廓特征进行建模。建模时,首先得到人体区域的最小竖直矩形,并提取矩形区域内的人体轮廓,定义矩形高度为 h ;然后从人体最高点开始,用 h 条水平线去截取人体轮廓,则每条水平线可以截得一左一右两个截点,于是可以得到以下两个特征向量: $\mathbf{R} = (r_1, r_2, \dots, r_h)$ 和 $\mathbf{L} = (l_1, l_2, \dots, l_h)$ 。其中 r_i 表示第 i 条水平线与人体轮廓截得的右交点的横坐标, l_i 表示第 i 条水平线与人体轮廓截得的左交点的横坐标,并且定义向量 $\mathbf{W} = (w_1, w_2, \dots, w_h)$, 其中 $w_i = r_i - l_i$ 表示第 i 条水平线处的人体宽度。特征向量 $\mathbf{R}, \mathbf{L}, \mathbf{W}$ 可分别拟合成函数 $f_R(x), f_L(x), f_W(x)$, 其中 $x \in [0, h]$, 当 x 为整数时,有 $f_R(x) = r_x, f_L(x) = l_x, f_W(x) = w_x$, 以下有关人体轮廓的应用都是通过这 3 个拟合函数实现的。

2.2 人体头肩的定位方法

针对直立行走的人体对象,本文采用了一种基于轮廓特征的人体头部定位方法。人体区域和轮廓如图 2 所示。根据 2.1 节的轮廓数学描述方式,由于人体头部轮廓呈现出椭圆特性,因此表示人体宽度的函数 $f_w(x)$ 在头部分界线处将出现局部极小值。图 3 是一条典型的人体宽度函数 $f_w(x)$ 曲线。

这里令 \mathbf{X}_{\min} 为函数 $f_w(x)$ 所有取得局部极小值的 x 值的集合,并规定 $x > 0$, 则人体头部分界线的位置定义为 $x_{HL} = \min(x_i), x_i \in \mathbf{X}$, 即集合 \mathbf{X} 中的最小值。考虑到人体头、肩位置存在一定的物理比例关系,本文将肩部位置定义为 $x_{SL} = x_{HL} + k \times x_{HL}$, 这里 k 为比例系数,需要通过实验确定,本文取 $k = 1/3$ 。图 4 为采用该方法对不同位姿下的人体进行头肩定位的示意图,图中在人体最小竖直矩形区域内有两条黑色的分割线,上面的分割线表示人体头部位置,下面的分割线表示肩部位置。需要指出的是,

这种基于人体轮廓特征的头肩分割方法对于自然情况下的直立人体对象具有良好的适应能力,且算法复杂度较低。当人的头部轮廓呈现非椭圆特性(例如特殊的发型,佩戴了帽子或其他修饰物)时,该方法将不再适用。此时可以采用现有的其他头肩检测的方法来定位人体头肩的位置。



图 2 人体头部分界线示意图

Fig. 2 Example of human head segmentation

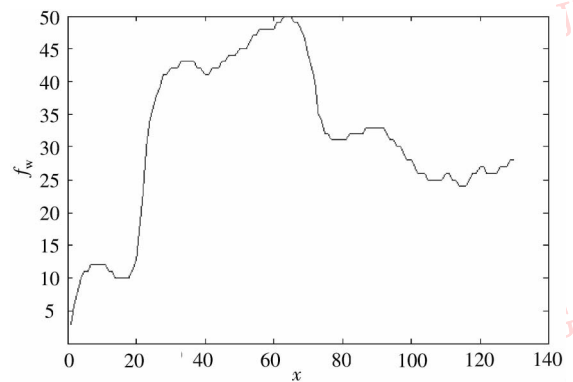


图 3 人体宽度 $f_w(x)$ 曲线

Fig. 3 A curve of the human body width $f_w(x)$

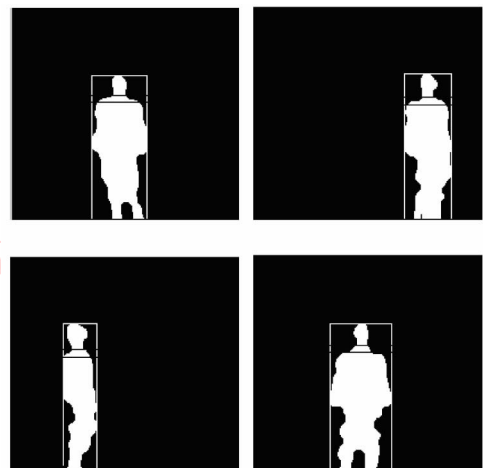


图 4 人体头肩定位示意图

Fig. 4 Example of head-shoulder localization

2.3 人体头部位姿的计算

假设人体头部区域面积用 S_H 表示,其大致可分为肤色区域和非肤色区域两部分。文献[12]给出了一种肤色模型,用于检测图像中的肤色像素,在已知的人体头部区域中,使用该文献的肤色检测方法便能得到肤色区域和非肤色区域。

图 5 给出了两幅图像,图像中人的头部呈现出不同的位姿。如果将图像中的头部区域沿其中心线分成左右对称的两块,那么当人脸面向或是背向摄像头时,左右两块区域的灰度分布是大致对称的(如图 5(a)所示),而当人脸面向摄像头的左侧或右侧时,左右两块区域的灰度分布存在明显的区别(如图 5(b)所示)。

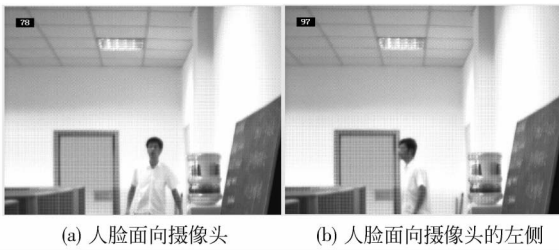


图 5 人体头部位姿示意图

Fig. 5 Example of human head pose

令 S_{LF} 、 S_{RF} 、 S_H 分别表示人体头部左半部分肤色区域的面积、人体头部右半部分肤色区域的面积以及整个头部区域的面积,则人体头部位姿变量 P_H 定义为

$$P_H = \begin{cases} 1 & \begin{cases} |S_{LF} - S_{RF}| < T \\ \frac{S_{LF} + S_{RF}}{S_H} < 0.5 \end{cases} \\ 0 & \begin{cases} |S_{LF} - S_{RF}| < T \\ \frac{S_{LF} + S_{RF}}{S_H} \geq 0.5 \end{cases} \\ \frac{S_{LF} - S_{RF}}{S_H} & |S_{LF} - S_{RF}| \geq T \end{cases} \quad (1)$$

式中, T 为事先设定的阈值, $P_H = 1$ 表示人脸背向摄

$$P_B = \begin{cases} 0 & |[f_L(x_{SL}) - f_L(x_{HL})] + [f_R(x_{SL}) - f_R(x_{HL})]| < T \\ \frac{f_L(x_{SL}) - f_L(x_{HL})}{x_{SL} - x_{HL}} + \frac{f_R(x_{SL}) - f_R(x_{HL})}{x_{SL} - x_{HL}} & |[f_L(x_{SL}) - f_L(x_{HL})] + [f_R(x_{SL}) - f_R(x_{HL})]| \geq T \end{cases} \quad (2)$$

式中, T 为事先设定的阈值, x_{HL} 表示人体头部分界线的位置, x_{SL} 表示肩部分界线的位置,显然有 $x_{SL} > x_{HL}$ 。式中的第 1 项 $\frac{f_L(x_{SL}) - f_L(x_{HL})}{x_{SL} - x_{HL}}$ 表示人体左轮

廓线, $P_H = 0$ 表示人脸面向摄像头,当 $P_H > 0$ 时,则表示人脸面向摄像头左侧, $P_H < 0$ 表示人脸面向摄像头右侧。式(1)表达的含义是首先计算 S_{LF} 和 S_{RF} ,当两者较接近时,则判定人脸面向摄像头或背向摄像头,在此基础上再计算 $\frac{S_{LF} + S_{RF}}{S_H}$,如果 $\frac{S_{LF} + S_{RF}}{S_H} > 0.5$,则头部区域中,肤色区域的比例较大,即判定人脸面向摄像头,否则,判定人脸背向摄像头;如果 S_{LF} 和 S_{RF} 差异较大,则可判定人脸面向摄像头的左侧或右侧。通过计算 $\frac{S_{LF} - S_{RF}}{S_H}$ 的值,就可以给出最终的判断,若该值大于 0,则表示人脸面向摄像头左侧,若该值小于 0,则表示人脸面向摄像头右侧。

2.4 人体躯干位姿的计算

令 $f_w(x_{HL})$ 表示人体肩部的宽度,笔者发现,随着人体侧身幅度的不断变大, $f_w(x_{SL})$ 逐渐变小,并且与摄像头较近的一段轮廓线在肩部区域的变化较快(如图 6 所示)。

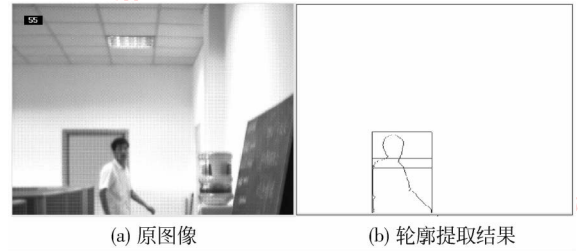


图 6 人体侧身时肩部区域的左右轮廓线

Fig. 6 The left and right contour of human shoulder when turning sideways

图 6(a) 为原图像,左侧的人体轮廓距摄像头较近,图 6(b) 为轮廓提取结果,并进行了头肩定位。显然,左侧轮廓线的变化较大,而右侧轮廓线的变化则相对平缓。利用这一人体特殊的成像规律,本文定义的人体躯干位姿变量 P_B 的计算方法如下:

廓线在肩部区域的梯度,通常为负值;第 2 项 $\frac{f_R(x_{SL}) - f_R(x_{HL})}{x_{SL} - x_{HL}}$ 表示右轮廓线在肩部区域的梯度,通常为正值。 $P_B < 0$ 表示左轮廓线在肩部区域的梯

度较大,左侧的人体轮廓离摄像头较近; $P_B > 0$,表示右侧的人体轮廓离摄像头较近; $P_B = 0$,表示左右轮廓线在肩部区域基本对称,此时人体面向或是背向摄像头。通常 P_B 的绝对值越大,人体的侧身幅度越大。

2.5 人体位姿所属状态的判别

在计算出了人体头部位姿变量 P_H 和躯干位姿变量 P_B 后,还需要通过一定的判别规则,才能估计最终的人体位姿。由于人在正常行走时,其头部位姿和躯干位姿存在一定的约束关系,根据经验,可以定义如下的判别规则:

- 1) $P_H = 0, P_B = 0$ 时,人体为状态 s_1 ,即面向摄像头;
- 2) $P_H = 1, P_B = 0$ 时,人体为状态 s_2 ,即背向摄像头;
- 3) $P_H \geq 0, P_B < 0$ 时,人体为状态 s_3 ,即面向摄像头向右侧身;
- 4) $P_H \leq 0, P_B > 0$ 时,人体为状态 s_4 ,即面向摄像头向左侧身;
- 5) $P_H < 0$ 或 $P_H = 1, P_B < 0$ 时,人体为状态 s_5 ,即背向摄像头向右侧身;
- 6) $P_H > 0$ 或 $P_H = 1, P_B > 0$ 时,人体为状态 s_6 ,背向摄像头向左侧身。

将得到的 P_H 和 P_B 的值代入以上规则就可以实现视频中的人体位姿估计。

3 仿真实验

为了检验算法的性能,本文通过 4 个实验与传统的 3 维重构方法进行比较。实验中使用的微机配置为 Intel(R) Pentium(R) 4, CPU 1.7GHz 256MB 内存,仿真实验采用 VC++ 编程。

实验 1 是使用 USB 网络摄像头采集的一段帧数为 500 帧的视频进行的实验,拍摄视频的光照为自然光,整段视频中存在光照的变化,图 7 中的 4 幅图像分别为实验视频中第 100 帧、200 帧、300 帧和 400 帧,其中图 7(c)(d) 两图的光照较强烈。采用本文的算法计算图像中的人体头部位姿 P_H 和躯干位姿 P_B ,并通过判别规则估计人体位姿。得到的实验结果如表 1 所示。



图 7 实验 1 实验视频中的 4 幅图像

Fig. 7 4 samples of test video

表 1 本文算法对图 7 进行人体位姿估计的结果

Tab. 1 The human pose estimation results of 4 samples in Fig. 7 using the proposed algorithm

	图 7			
	(a)	(b)	(c)	(d)
P_H	1	0	0.12	-0.1
P_B	0	0	1.5	-1.9
位姿姿态	s_2	s_1	s_3	s_4

实验 2 是使用带图像采集卡的专业摄像头采集的一段帧数为 200 帧的视频(见图 8)进行的实验。摄像时用的光照为室内灯光。位姿估计的结果如表 2 所示。



图 8 实验 2 实验视频中的 4 幅图像

Fig. 8 4 samples of test video in experiment 2

表 2 本文算法对图 8 进行人体位姿估计的结果

Tab. 2 The human pose estimation results of 4 samples in Fig. 8 using the proposed algorithm

		图 8			
		(a)	(b)	(c)	(d)
P_H		0	0	0	0.3
P_B		-1.8	-1.7	0	2.1
位姿姿态		s_4	s_4	s_1	s_3

实验结果与实际情况吻合。与实验 1 的实验效果类似,实验视频的 200 帧图像中检测正确的图像为 178 帧,正确率接近 90%,处理速度为 20 帧/s。实验 3 是对在室外的自然环境下采集的一段帧数为 200 帧的视频进行的实验。场景中背景环境相对复杂,视频中同时存在两个运动人体。图 9 选取了视频中两幅有代表性的图像来说明实验结果。



图 9 实验 3 实验视频中的 2 幅图像

Fig. 9 2 samples of test video in experiment 3

采用本文算法分别对图中两个运动人体的位姿进行计算。对于图 9(a)黑框中的人体,其 $P_H = 0.5, P_B = 0.7$,位姿估计为面向摄像头向右侧身;对于图 9(b)黑框中的人体,其 $P_H = 0.3, P_B = -1.3$,位姿估计为背向摄像头向左侧身。实际情况证明,算法对黑框中人体的位姿估计是正确的。而两图中黑框外的人体位姿估计则存在误差,估计结果均为面向摄像头,与实际情况不符。

在上面 3 个实验中,出现位姿估计错误的原因主要有以下两点:1)由于图像中的噪声干扰或是摄像机的突然抖动等造成人体区域的检测效果不理想,导致人体轮廓的提取不够精确,从而影响了最后的位姿估计结果;2)由于人体在运动中出现低头、手臂摆动幅度太大等不规范的行走姿态(如实验 3 中黑框外的人体)所致。

实验 4 是采用传统的 3 维重构方法,对实验 1 中的视频进行人体位姿估计。实验中先选取人体肩部位置的左右两端点作为关键特征点(如图 10 所

示),然后利用 3 维重构的方法计算关键特征点的深度值,即其到摄像头的距离,单位为 m,对于图 7 中的 4 幅图像,得到的实验结果如表 3 所示。

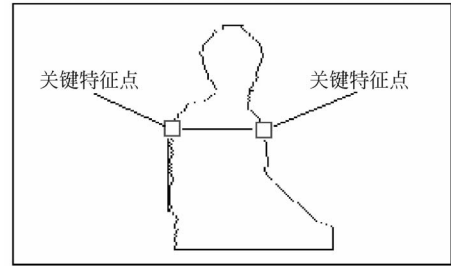


图 10 关键点选取示意图

Fig. 10 Example of key points selection

表 3 利用传统的 3 维重构方法得到的实验结果

Tab. 3 The test results using 3D reconstruction method

		图 7			
		(a)	(b)	(c)	(d)
左端点深度/m		4.67	4.45	5.40	5.11
右端点深度/m		4.75	4.40	5.06	5.34

从上面的结果可以看出,3 维重构的方法能够从 2 维图像中得到特征点的深度值,进而估计人体的侧身角度。例如表 3 中的图 7(a)(b)两幅图像中的左、右特征点的深度值大小接近,表示人体侧身角度不大,图 7(c)(d)两图中左、右特征点的深度值存在一定的差异,表示人体侧身角度较大。然而该方法无法判断图像中的人体投影是正面还是背面,另外,当目标区域检测不准确时,其计算结果也会出现错误。利用 3 维重构的方法处理视频中的其他图像时,可以得到相似的效果,处理速度为 6 frame/s,可见其效率不如本文算法。

通过前 3 个实验可以看到,本文算法在不同成像器件、不同场景以及光照变化下,当位姿估计对象为正常行走时,实验结果具有良好的稳定性和有效性,而通过实验 4 可以给出本文算法与传统 3 维重构方法的性能比较(如表 4 所示)。在处理速度上,本文算法为 20 帧/s,传统的 3 维重构方法为 6 帧/s,可见本文算法的计算复杂度较小;在实际应用时本文算法仅需单摄像头就能完成,而 3 维重构方法则需要多个摄像头,并且实验前需要测量各摄像头的光学参数和位置参数;在位姿估计效果上,本文算法虽不能得到人体位姿的精确量化值,但能对人体位姿进行合理的估计,并能给予自然语言解释,而 3 维

重构方法虽然能估计人体侧身的角度,却无法识别图像中的人体投影是正面投影还是背面投影。

表 4 本文算法与传统的 3 维重构方法的性能对比

Tab.4 Comparison of the performance of proposed algorithm and 3D reconstruction method

	算法效率	应用限制	位姿识别效果
本文算法	较快	少	能对人体位姿进行合理的估计,并能给予自然语言解释
3 维重构	较慢	多	只能估计人体侧身角度,无法判断朝向

4 结 论

传统的 3 维重构方法在计算人体位姿时,计算复杂度高,应用限制多。本文针对这一问题提出了一种基于头肩分割的人体位姿估计算法。该算法首先将人体位姿简化为有限状态空间,然后利用人体在 2 维成像时的规律和特点来估计人体位姿,并给予合理的自然语言解释。该算法与传统的基于 3 维重构的方法相比,不需要预先标定摄像头参数,仅利用单目视觉信息就能快速有效地获得人体位姿信息。实时性和应用便捷性是本文算法的主要特点。当前该算法的适用对象为无遮挡的直立行走人体,在有遮挡或人体处于其他姿态时(跑、跳、蹲等)如何估计人体位姿是下一步的研究工作。

参考文献 (References)

- [1] Wang Liang, Hu Weiming, Tan Tielu. Recent developments in human motion analysis[J]. Pattern Recognition, 2003, 36(3): 585-601.
- [2] Bobick A F, Davis J W. The recognition of human movement using temporal templates [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 15(3): 257-267.
- [3] Luo Y, Wu T D, Hwang J N. Object-based analysis and interpretation of human motion in sports video sequences by dynamic Bayesian networks [J]. Computer Vision and Image Understanding, 2003, 92 (2): 196-216.
- [4] Oliver N, Horvitz E. A comparison of HMMs and dynamic Bayesian networks for recognizing office activities[C]// Lecture Notes in Artificial Intelligence. Berlin, Germany: Springer Verlag, 2005, 3538: 199-209.
- [5] Kolonias I, Christmas W, Kittler J. Use of context in automatic annotation of sports videos [C]// Lecture Notes in Computer Science. Berlin, Germany: Springer Verlag, 2004, 3287: 1-12.
- [6] Cho K, Cho H, Um K. Human action recognition by inference of stochastic regular grammars [C]// Lecture Notes in Computer Science. Berlin, Germany: Springer Verlag, 2004, 3138: 388-396.
- [7] Yamato J, Ohya J, Ishii K. Recognizing human action in time-sequential images using Hidden Markov model[C] //Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Piscataway, NJ, USA: IEEE Computer Society, 1992, 379-385.
- [8] Arie J B, Wang Z, Pandit P, et al. Human activity recognition using multidimensional indexing [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(8): 1091-1104.
- [9] Park S, Subbarao M. Automatic 3D model reconstruction based on novel pose estimation and integration techniques[J]. Image and Vision Computing, 2004, 22(8): 623-635.
- [10] Kazuhiko Takahashi, Yusuke Nagasawa, Masafumi Hashimoto. Remarks on 3D human posture estimation system using simple multi-camera system [C]//Proceedings of IEEE International Conference on Systems, Man, and Cybernetics. Piscataway, NJ, USA: IEEE Computer Society, 2006: 1962-1967.
- [11] Zhou Jianpeng, Hoang Jack. Real time robust human detection and tracking system [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE Computer Society, 2005: 149-156.
- [12] Karungaru S, Fukumi M, Akamatsu N. Detection of human faces in visual scenes[C]//Proceedings of the Seventh Australian and New Zealand Intelligent Information Systems Conference. Brisbane, Australia QLD: The Australian Pattern Recognition Society, 2001: 165-170.