

基于多约束融合的人手臂三维运动分析

潘春洪 马颂德

(中国科学院自动化研究所模式识别国家重点实验室, 北京 100080)

摘要 为了在没有任何特殊标志的情况下, 实现从单目序列图象中分析、估计人手臂的二维运动, 提出了一种多约束融合的方法, 该方法是利用棍棒模型来模拟人的手臂, 首先通过处理单目图象序列来自动获取图象序列中手臂关节点的对应; 然后再利用多约束融合及基于图象序列中关节点的对应, 即估计尺度意义下关节点的二维相对运动轨迹; 最后利用真实图象来获得相应人手臂的三维运动轨迹, 并将其与通过运动捕捉系统获得的人手臂的真实三维运动轨迹进行了比较实验. 实验结果表明, 该方法用于对人手臂的运动分析非常有效.

关键词 多约束融合 运动分析 平滑运动

中图分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2006)11-1053-05

3D Motion Analysis of Human Arm Based on Multi-Constraints Fusion

PAN Chun-hong, MA Song-de

National Laboratory of Pattern Recognition, Institute of Automation Chinese Academy of Sciences, Beijing 100080

Abstract Motion estimation of human from a single or a monocular sequence of image is an important research field in Computer Vision and Image Processing. In this paper, based on constraint fusion, we proposed 3D-motion estimation from the monocular perspective images without any special markers. We consider the arm as stick model with three joints, and the link between every two neighboring joint points is regarded as rigidity. Based on the knowledge of anatomy and kinematics, we exert coplanar constraint, rigidity constraint, smooth motion constraint on this model. By pre-processing the image sequence, we semi-automatically yielded 2D correspondences of joint points of arm among images. Then, by constraint fusion, 3D relative structure of arm can be determined in a scale factor. The experiment with real images is included. In order to testify the validity of the proposed method, we compare the 3D motion of human arm obtained by the motion capture system with that obtained by our method. The result shows that the method is very efficient.

Keywords Constraint fusion, Motion analysis, Smooth motion

0 引言

众所周知, 由于现有的人机交互设备, 如键盘、鼠标已经成为有效利用信息流的瓶颈, 特别是随着新的显示技术, 如虚拟现实技术的出现, 一些其他的人机交互方式, 如语音交互、手势交互, 将会取代现有的键盘和鼠标, 所以, 作为一种非语言的人机交互方式, 人的手臂(或手)的姿势识别在人机交互领域

将会越来越显示出其重要性.

过去, 人们曾试图利用简单的背景或对用户施加一些约束(如手套、标志等)来捕捉人的三维运动, 并由此产生了所谓的数据手套及运动捕捉系统. 虽然这些设备能获得高精度的三维信息, 但是使用者不得不穿着一些笨重的设备, 或在关节点上放一些特殊的标志点, 同时往往还需要戴着一些连接设备和计算机之间的缆线, 因而这大大影响了人机交互, 也限制了这些设备或系统的应用范围. 如果人的三

基金项目: 中国科学院自动化所与法国 IMAGES 的合作项目“3D Geometric and Photometric Modeling of the Scene for Digital Museum”的主要研究内容之

收稿日期: 2006-03-20; **改回日期**: 2006-12-12

维运动跟踪能够更加自动,更加廉价,并更加实用和快捷,那么许多应用,如动画、虚拟现实的人机交互将更加容易实现。

虽然许多文献均论述了人手臂的运动跟踪系统,文献[1]还给出最近几年这方面工作的综述,如 Hoffman 等人首先提出了将人的解剖约束应用于人的运动分析中^[2],然而这些方法大多是基于回归滤波的方法,如卡尔曼滤波;之后 Azoz 等人强调了多模型方法^[3],他们把多种约束融合在扩展的卡尔曼滤波器中,以用于估计人手臂的三维运动,但是由于这种方法过分依赖于状态变量的选择,如果变量选择不适当,那么被估计的运动和真实运动之间就会有很大的差异,因此,Hiroyuki 等人利用基于扭矩的动力学系统来获取人手臂的三维运动^[4],事实上,由于物体运动的惯性,一个物体的运动状态不可能瞬间改变,因此,在运动分析中,运动的平滑性约束是一个非常合理的假设,然而,到目前为止,还没有人将物体运动的平滑性约束用于人手臂运动的三维估计中。

本文提出了一种用于估计人手臂三维运动的多约束融合方法,该方法是首先将人的手臂看作一个带有一些关节点(肩、肘、腕)的棍棒模型;然后通过处理序列图象来获取和跟踪这些关节点;最后利用刚体约束、结构约束、平滑约束等多约束融合的方法来估计人手臂的三维运动轨迹,为了测试该方法的有效性,本文还用运动捕捉系统获取了一些人手臂的真实运动轨迹,并将其与利用本文所述方法所获取的数据进行了比较实验,实验结果表明,该方法行之有效。

1 手臂的棍棒模型

如前所述,本文将手臂看作是带有3个关节点(肩、肘、腕)的棍棒模型,且关节点之间的连接体是刚体,为了使该模型尽可能简单,并具有尽可能少的自由度,本文没有考虑手臂的形状,并从解剖学和运动学的角度对该模型施加如下约束,即首先两关节点之间的距离不变,这是运动中的刚体约束;更进一步地,根据解剖学的知识可知,手臂长度和人的身高具有一定的比例关系,比如,前臂和后臂分别是人身高的0.206和0.22倍;另外,前后臂之间的角度也有一定约束,本文的方法将包含所有这些约束。

2 关节点的提取和对应

在关节点提取时,首先用梯度水线运动分割方法和基于 Voronoi 图的骨架提取方法,来获取序列图象中人的骨架(见图1~图3),详见文献[5],接下来再基于获取的骨架,从序列图象中提取手臂的关节点(肩、肘、腕),由此即可获得序列图象中关节点的对应,为了获得这些关节点,首先需要参数化骨架,通常,因为手臂骨架的自动获取是非常困难的,所以这里,需首先从第一帧图象中用人工方法来获取手臂的肩关节点,然后利用相邻图象间灰度相关的原理来获取该点在序列图象中的相应位置,并考虑手臂骨架的几何信息,来找到序列图象中肩关节的对应。

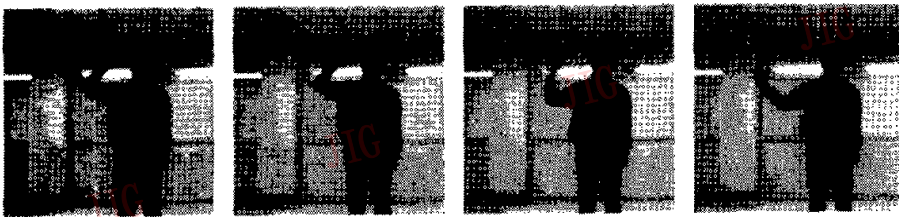


图1 用于采样的灰度图象序列



图2 分割后的二值边缘图象



图 3 带有被跟踪特征点的骨架图象

一旦获取了肩关节的对应,就能从序列图象中自动获取手臂的骨架,然后应用自然三次样条(见下式)将手臂的骨架参数化,这里样条平滑的近似判决条件是如下最小二乘估计:

$$x(u) = \sum_{k=1}^n (g_k \cdot BF_k(u))$$

式中, g_k 是骨架的几何约束参数; $BF_k(u)$ 是多项式混合函数; $x(u)$ 代表骨架上任一点的坐标, $0 \leq u \leq 1$.

最后,本文用参数化骨架的几何信息和手臂的边缘信息去获取另外的关节坐标,其中,骨架的延长线与手臂边缘的交点,可认为是腕关节,而参数化骨架的最大曲率点则是肘关节,利用上述方法所获取的关节的相对误差大约为 15% 左右.

3 基于多约束融合的三维运动估计

运动分析和刚体约束:第 i 个特征点 $P_i(t)$ 在 t 时刻的三维空间坐标能被表示成 $P = (X, Y, Z)$, 而将它 在 图 象 上 的 投 影 表 示 成 $P = (x, y)$, 这里,下标 i 表示第 i 个特征点,其上标 j 表示采样时刻 $t_j (z_j^{(i)} = z_j^{(i)}, z_j^{(i)} = z_j^{(i)})$, 而且当使用针孔摄像机模型时,三维坐标和图象坐标之间的关系能被表示成

$$x_j = \frac{fX_j}{Z_j}, \quad y_j = \frac{fY_j}{Z_j}$$

式中, f 是摄像机焦距.

本文就是用上面的方法处理了手臂的序列运动图象,同时获取了关节点在序列图象中的对应,这样就能获取关节点的图象坐标,其目的就是通过各关节点的图象坐标的对应,再利用多约束融合,去恢复各关节点的三维坐标.

由于肩关节的运动相对肘关节和腕关节的运动是很小的,因此,这里可假设肩关节的深度坐标是相对固定的.基于此假设,虽然肩关节的图象坐标是变化的,但是, Z_j 却保持不变,而且可将手臂分成上

臂和下臂两部分,其每一部分包括一个单刚体.由于本文假设肩关节的深度坐标不变,因此,可首先分析比较简单的上臂运动情况,并确定上臂的三维结构,由此就能分析下臂的三维运动,接着利用刚体约束,就可以得到上臂的刚体约束方程

$$\|P_j^{(j)} - P_j^{(j)}\| = l; \quad j = 0, 1, 2, \dots$$

这里, $P_j^{(j)}, P_j^{(j)}$ 分别是肩关节和肘关节的三维坐标, l 是上臂的长度.

由于肩关节的深度坐标固定,由上述方程则可得到

$$(x_j^{(j)2} + y_j^{(j)2} - f^2)Z_j^{(j)} - 2(x_j^{(j)}x_j^{(j)} + y_j^{(j)}y_j^{(j)} + f^2)Z_j^{(j)}Z_j^{(j)} - (x_j^{(j)2} + y_j^{(j)2} + f^2)Z_j^{(j)2} = (lf)^2$$

式中, $Z_j, Z_j^{(j)}$ 和 l 是未知变量,但事实上,如果肩关节的深度坐标固定,那么, Z_j 就仅仅是一个尺度,因此,对于给定的 N 幅图象,在尺度意义下就有 $N+1$ 个变量, N 个方程,为了使方程组可解,必须增加更多的约束.

3.1 结构约束

根据解剖学的知识^[1]可知,连接两部分之间的关节有着一定的自由度,比如肩关节有 3 个自由度,而肘关节只有一个.由此可见,对于大部分关节而言,其运动范围是确定的,而对于肘关节来说,其运动范围必须是 $20^\circ \leq \theta \leq 180^\circ$. 另外,虽然肩关节有 3 个自由度,但因其旋转角并非任意的,所以,对于手臂的三维运动估计,还可增加一些结构约束,当 l, Z_j 给定时,上述刚体约束方程是一个关于未知变量 $Z_j^{(j)}, Z_j^{(j)}$ 的二次方程,其解必须是物理上可实现的,即二次方程的判别式应该大于零.这样,当给定某一范围时,就能获取一组肘关节的运动轨迹,而且一旦求出肘关节的运动轨迹,就可以利用上下臂之间的长度约束来求出腕关节相应的三维运动轨迹,但这组运动轨迹必须满足手臂运动的结构约束.事实上,手臂的运动轨迹可以被看作是一组关于时域和空域参数空间的任意曲线,而手臂的运动识别其实就是增加参数空间的约束条件,以获取其真正的运动轨迹.

3.2 运动的平滑性约束

众所周知,由于惯性,一个物体的运动不可能产生突变,即如果序列图象以一定的采样间隔采样(本文的实际采样间隔为 1/10s),那么每相邻两帧图象之间就不会有运动突变,因此运动分析中,运动的平滑性假设是一个合理的假设.

假设 $\hat{\xi}_i$ 表示第 k 帧肘关节的运动轨迹的偏差,其表达式如下

$$\hat{\xi}_i = \varphi \overline{P_{i-1}^{(k)} P_i^{(k)}} \cdot \overline{P_{i-1}^{(k)} P_i^{(k)'}}$$

由于该式考虑了序列图象中物体在每帧中位移的大小和方向,因此,又可以定义运动轨迹的偏差为

$$J = \sum_{i=1}^{N-1} \hat{\xi}_i$$

运动的平滑性意味着该轨迹的偏差 J 有最小值,这里,由于可用两帧间物体运动的距离和角度的变化作为轨迹的偏差,因此上式又可以表示成

$$J_1 = \min \sum_{j=1}^{N-1} |d_j - \bar{d}|, J_2 = \min \sum_{j=1}^{N-1} |\theta_j - \bar{\theta}|$$

$$\text{式中, } \bar{d} = \frac{1}{(N-1)} \sum_{j=1}^{N-1} d_j, \bar{\theta} = \frac{1}{(N-2)} \sum_{j=1}^{N-2} \theta_j,$$

其中, d_j 是肘关节特征点从 $P_{j-1}^{(k)}$ 运动到 $P_j^{(k)}$ 之间的距离 ($j=1, 2, \dots, N-1$), 角度 θ_j 是直线 $\overline{P_{j-1}^{(k)} P_j^{(k)'}}$ 和 $\overline{P_j^{(k)} P_{j+1}^{(k)'}}$ 之间的夹角 ($j=1, 2, \dots, N-2$).

本文所述的算法是用于解决一个多约束融合优化问题的,在处理该问题时,首先用刚体约束方程来获取一组 $l-2$ 参数空间的肘关节运动轨迹;然后再利用刚体约束和上下臂之间的长度关系来获取一簇与肘关节运动相应的腕关节运动轨迹,该轨迹满足上述的结构约束,这样,就可以得到一簇手臂的三维运动轨迹,且该簇运动轨迹均满足刚体约束和结构约束;最后,运用运动的平滑性约束,将从该簇运动

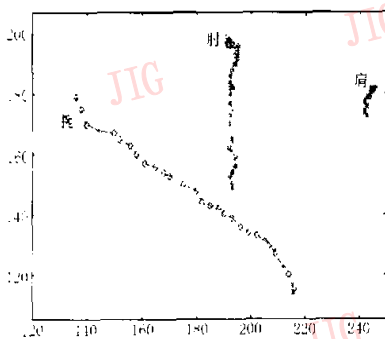
轨迹中找出的一条由本文所定义的最平滑的轨迹作为所求手臂的三维运动轨迹.

4 实验结果

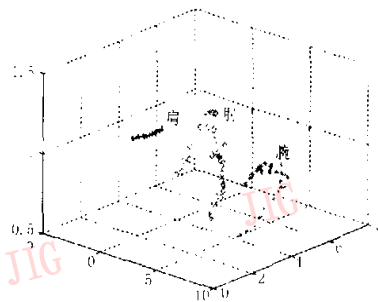
本文用真实图象来进行测试实验.测试时,首先用已标了的摄像机来获取一组带有手臂运动的序列图象(图 1),图 1 即为一组带有手臂运动的序列图象中的一些采样帧;然后用上述基于梯度水线的图象分割方法和基于 Voronoi 图的骨架提取方法,来预处理该序列图象;最后通过参数化手臂的骨架,再利用序列图象的灰度相关和参数化骨架的几何信息来获取关节点在序列图象中的对应运动轨迹.图 2 和图 3 分别为手臂运动图象分割后的二值边缘图象和带有有关节点对应的该序列图象的骨架图象.图 4(a) 显示了肩关节、肘关节和腕关节图象运动轨迹.图 4(b) 为相应关节点的三维运动轨迹.

通过用上述多约束融合的方法以及序列图象中手臂各关节点的对应,就可以求出尺度意义下肩关节、肘关节和腕关节的三维运动轨迹.这也就是说,一旦上臂或下臂的真实长度确定后,就能得到真实肩关节、肘关节和腕关节的三维运动轨迹.

为了测试本文方法的有效性,还用精度高达 0.1% 的运动捕捉系统获取了一组人手臂的真实运动数据,图 5(a) 即为一幅带有人手臂运动的捕捉系统场景图象,同时在人手臂的关节点处放置了一些特殊的标志点,并用摄像机来获取一组序列图象;然后通过跟踪序列图象中手臂上的 3 个关节点来获取与该序列图象上关节点对应的运动轨迹(图 4(a));最后用本文提出的多约束融合的方法获得了一组这些关节点的三维运动轨迹(图 4(b)).当然,因为运



(a) 跟踪获得的图象上肩、肘和腕关节运动轨迹

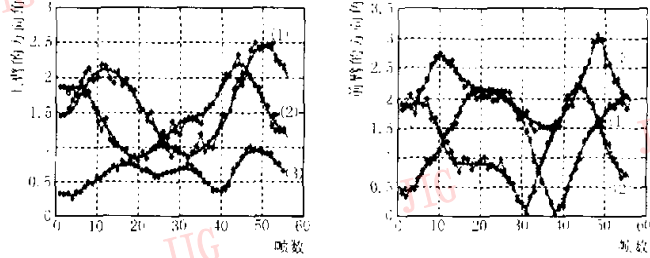


(b) 肩、肘和腕关节点的三维运动轨迹

图 4 用本文方法获得的肩、肘、腕运动轨迹



(a) 关节处放置了标志点的手臂运动图象



(b) 肩关节、肘关节和腕关节在不同时刻的位置矢量

(其中带有·的曲线是通过运动捕捉系统获取的,带有△的曲线是通过本文方法获取的)
(1) 肩关节;(2) 肘关节;(3) 腕关节

图 3

动捕捉系统获取的是手臂真实的三维运动轨迹,而本文估计的是手臂的相对运动轨迹,所以直接将这两组运动轨迹进行比较是毫无意义的。这里,可通过计算和归一化两组运动手臂在不同时刻的位置矢量,图 3(b)给出了肩关节、肘关节和腕关节在不同时刻通过两种方法获取的手臂位置矢量。虽然两者之间的相对误差高达 20% 左右,但从位置矢量曲线总体上看,估计的手臂运动轨迹和真实的运动轨迹是非常相似的。

在比较实验中,还发现摄像机和手臂的相对位置对手臂的三维跟踪是非常敏感的,特别当手臂的运动垂直于摄像机平面时,由于其被估计的运动参数非常不准确,也即此时所获得的数据是不可靠的,是退化了的,因此,为了增加运动恢复的鲁棒性,此时用带有约束融合的扩展卡尔曼滤波方法更为合理。

5 结 论

综上所述,本文提出了一种基于多约束融合的用于获取单目序列图象关节点对应的三维运动分析方法。在此基础上,又做了相关的实验,即首先通过对单目序列图象进行预处理来获取手臂关节点图象坐标的对应,然后利用刚体约束、结构约束和平滑运动模型去估计人手臂的三维运动,并将其与真实运动相比。实验结果表明,该方法用于手臂的三维运动分析是行之有效的。

参 考 文 献

1 Pavlovic V. I., Sharma R., Huang T. S. Visual interpretation of hand gestures from human-computer interaction: a review[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1997, 19(7): 677~695.

2 Hoffman D. D., Flinchbaug B. E. The interpretation of biological motion[J]. Biological Cybernetics, 1982, 42(7): 193~203.

3 Azou Y., Devi L., Sharma R. Tracking hand dynamics in unconstrained environment [A]. In: Proceedings of the International Conference on Computer Vision [C], Bombay, India, 1998; 274~279.

4 Hirokyi Segawa, Takashi Totsuka. Torque-based recursive filtering approach to the recovery of 3D articulated motion from image sequence [A]. In: Proceedings of Computer Vision and Pattern Recognition [C], Fort Collins, Colorado USA, 1999; 340~345.

5 Chunhong PAN, Songde MA. Motion capture of arm from a monocular image sequence [A]. In: Proceedings of VISUAL [C], Amsterdam, 1999.



潘春洪 1987 年获清华大学自动化系, 1990 年获中国科学院上海光机所硕士学位, 2000 年获中科院自动化所模式识别与智能系统专业博士学位, 现为美国南加州大学计算机系博士后, 主要研究领域为计算机视觉、计算机图形学、模式识别。



马颂德 博士生导师, IEEE 高级会员, 1968 年毕业于清华大学自动化系, 1986 年获法国国家博士学位, 曾任中科院自动化所所长, 现为国家科技部副部长, 长期从事计算机视觉、图象处理和模式识别等方面的研究工作。