

一种人头部实时跟踪方法

徐一华 朱玉文 贾云得

(北京理工大学计算机科学与工程系, 北京 100081)

摘要 为了能够在视频监控、人机交互、视频会议等领域对人头部运动实施实时跟踪,给出了一种使用黑白摄像机对人平移或转身时的头部运动进行实时跟踪的方法.该方法主要由基于块特征的跟踪和基于头部几何特征的校正两个步骤组成.块特征跟踪算法仅利用图象低层信息而不依赖于目标的具体模型,可实现对头部自由运动的跟踪.为解决块特征跟踪误差累积等原因造成的目标丢失问题,又采用了头部轮廓几何特征检验方法,根据跟踪窗口中头部轮廓位置的偏移来对块特征跟踪结果进行校正.另外,为提高转身运动时相邻两帧图象的特征跟踪正确率,还引入圆柱模型来拟合头部,并在展开柱面内进行块特征选取和跟踪.本文方法在PⅡ 350微机进行了实验,实验结果表明,系统能对长时间图象序列中人平移或转身时头部运动实施准确跟踪.当跟踪窗口大小为 120×180 pixels,块特征数目为80个时,系统的处理速度达到30帧/s.

关键词 图象运动估计 头部跟踪 块特征 跟踪校正

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2002)01-0011-05

Real-Time Tracking of Head Motion

XU Yi-hua, ZHU Yu-wen, JIA Yun-de

(Department of Computer Science and Engineering, Beijing Institute of Technology, Beijing 100081)

Abstract This paper presents an improved real-time head tracking method, which can realize robust tracking of a person's translation and turning around with a black/white camera. This method consists of two main steps: block feature based head tracking, and head geometry based correction. The block feature based tracking only uses low-level image information and does not rely on the models of different objects, thus it can be used to track the free motion of the head. The correction step is introduced to address the shift caused by errors accumulating of block tracking. This step measures the displacements of the head contour in the tracking window of current image, and hereby to correct the results of block tracking. Moreover, to improve the performance of block tracking during person's turning around, we introduce a cylindrical model to approximate head, and extract and track features in the warped cylindrical surface. The resulting system provides robust and precise tracking over long sequence on a 350 megahertz microcomputer, and operates at 30 frames per second with the tracking window size of 120×180 pixels and the tracking sets of 80 features.

Keywords Image motion estimation, Head tracking, Block feature, Tracking correction

0 引言

人头部实时跟踪在视频监控、人机交互、视频会议等各种场合有着广泛的应用.人实施头部跟踪可利用的信息主要有几何信息、颜色信息和运动信息3种,其中,几何信息有头部形状、人脸轮廓、五官的形状及相

对位置等;颜色信息指脸部皮肤、头发的颜色及其分布等;运动信息则指头部表面各点之间的运动相关性以及整体的运动轨迹.结合使用这几种信息,现有的人头部跟踪方法大致可分为点特征方法、基于模型的方法、灰度分布法、肤色方法和块特征方法5类.其中点特征方法是通过提取眼角、鼻翼、嘴角等特征来实现人脸跟踪^[1].基于模型的方法则是通过对头部及其器官进行

建模,并利用模板匹配来实现跟踪^[2].灰度分布法是先求解三维空间内头部表面的灰度分布马赛克图,然后通过匹配来实现跟踪^[3].肤色方法则利用人脸肤色的特殊性来进行跟踪^[4,5],而实际系统常结合使用上述多种方法^[6],并用三维线框^[2]、圆柱模型^[3]、或椭圆模型^[7]等来拟合头/脸部的轮廓,以及卡尔曼滤波等进行运动预测^[5,6],以求获得比较鲁棒的解,但由于人头后部没有器官、肤色及足够的纹理信息可以利用,因此这些方法均难以实现完全转身情况下的正确跟踪.

基于块特征的方法是通过图象中批量块特征的运动矢量计算来实现目标跟踪^[8,9].它虽具有精度高、不依赖于具体目标模型等优点,可用于实现人头自由运动的跟踪,但由于块特征仅利用了图象的低层信息,且不能根据目标的整体形状来对跟踪结果进行调整,因此在长时间连续跟踪时,容易因误差累积而发生目标丢失的情况;同时在跟踪过程中(尤其头部旋转运动时)块特征需要不断更新,而累积误差则会造成更新的某些特征是非头部特征,这也会导致目标逐渐丢失.另外,人在转身运动时,头部表面各点和摄像机的角度关系会发生变化,同一特征块在不同帧中会有较大形变,因而也会使得转身时块特征会大量丢失.

本文以块特征跟踪为基础,结合利用头部轮廓几何特征,提出一种新的人头部运动跟踪方法.由于该方法可根据头部轮廓在当前跟踪窗口中的位置偏移来对块特征跟踪结果进行校正,因而可以实现长时间图象序列中人平移或转身运动的准确跟踪.为提高转身情况下块特征的跟踪率,本文引入圆柱模型来拟合人头部,并在展开后的圆柱表面内进行特征选取和跟踪.系统的主要组成如图1所示.

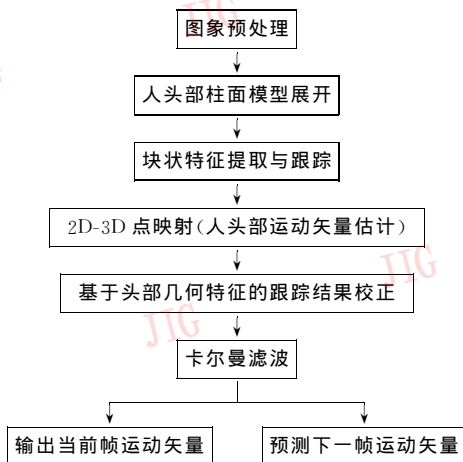


图1 人头部运动实时跟踪系统模型

1 基于块特征的人头部运动跟踪

1.1 图象块特征的跟踪

在目标运动图象序列中,如果采样频率足够高,则可假定相邻两帧图象中足够小的块状特征发生了一定位移,而其灰度值保持不变

$$I(x, t) = I(\delta(x), t + \tau) \quad (1)$$

其中, x 为特征初始位置, $\delta(\cdot)$ 为运动场, τ 为运动时间. 若以平移运动逼近其运动场,则有 $\delta(x) = x + d$, 其中, d 为位移向量,而跟踪算法的目的,即是为了计算连续帧间块状特征的位移值 d .

由于噪声等因素的影响,因而实际上特征窗口的灰度值变化不会完全为零. 上述问题可转化为求解一合适 d 值,并使之满足窗口间 SSD(the sum of squared difference) 值为最小的问题

$$E = \sum_{x, y \in W} w(x) [I(x + d, t + \tau) - I(x, t)]^2 \quad (2)$$

其中, W 为选定大小的特征窗口, $w(x)$ 是权值函数. 用一阶台劳级数逼近 $I(x + d, t + \tau)$, 并令权值函数为 1, 求解上式关于 d 的导数为 0, 得到如下线性方程组

$$Gd = e \quad (3)$$

其中

$$G = \sum_W \begin{bmatrix} I_u^2 & I_u I_v \\ I_u I_v & I_v^2 \end{bmatrix}, e = -\tau \sum_W I_t \begin{bmatrix} I_u \\ I_v \end{bmatrix}, d = \begin{bmatrix} d_u \\ d_v \end{bmatrix}$$

其中, $I_u = \partial I / \partial u$, $I_v = \partial I / \partial v$, $I_t = \partial I / \partial t$, 下标 u, v 表示图象的行列坐标轴. 当 G 为可逆矩阵时,即可解得特征运动向量的估计值, $d = G^{-1}e$.

由上式可知,只有当 G 为可逆矩阵时,运动向量 d 才是可解的,而且由于噪声等因素的影响,若该矩阵的某个特征值过小,则求解得到的运动向量也可能很不可靠,因此提取块特征时,可以以其 G 矩阵的较小特征值为标准. 若令该矩阵两个特征值分别为 λ_1, λ_2 , 则特征选取标准可定义为: $\min(\lambda_1, \lambda_2) > \lambda$, 其中, λ 为预定义的阈值. 在进行块特征跟踪时,使用多层分辨率模型,不仅可以提高算法运行速度,并可增大块特征的跟踪范围. 详细描述参见文献^[8,9].

1.2 人头部圆柱模型

由于和摄像机的角度关系不同,因而人头部表面上某些位置在图象中可能会产生严重的透视缩影效应. 这不仅使得转身时头部表面同一区域在摄像机中的成象产生很大差异,而且还导致块特征大量丢失.

Shi 等提出的特征仿射运动模型^[8]虽可在一定程度上解决这一问题,但该方法计算量很大,而且容易导致误匹配.本文引入圆柱模型来拟合人头部,并在展开后的圆柱表面内进行特征选取和跟踪,实验结果表明,这样可以提高转身时头部特征的正确跟踪率.

根据小孔成像模型,空间点坐标与其在图象中投影点坐标的关系如下:

$$\begin{Bmatrix} u_i \\ v_i \end{Bmatrix} = f \cdot \begin{Bmatrix} x_i/z_i \\ y_i/z_i \end{Bmatrix} \quad (4)$$

其中, x_i, y_i, z_i 为空间点 i 在摄像机坐标系中的坐标, u_i, v_i 为其投影点在图象坐标系中的坐标, f 为摄像机焦距.假定空间点深度值 z_i 变化范围不大,并选取合适的空间坐标系以规一化比例因子 f/z_i , 则该模型可简化为如下正射投影模型

$$\begin{Bmatrix} u_i \\ v_i \end{Bmatrix} = \begin{Bmatrix} x_i \\ y_i \end{Bmatrix} \quad (5)$$

采用圆柱模型来拟合人头部,并使用该正射投影模型,可得到如下头部表面在图象中投影的展开公式

$$\begin{cases} u_{r,i} = \cos\alpha \cdot (r \cdot \arcsin(t/r) - t) + u_i \\ v_{r,i} = \sin\alpha \cdot (r \cdot \arcsin(t/r) - t) + v_i \end{cases} \quad (6)$$

$$t = \frac{\cos\alpha(u_i - u_0) - \sin\alpha(v_i - v_0)}{\cos^2\alpha - \sin^2\alpha}$$

其中, u_0, v_0 为人头部圆柱中心点坐标, α 为圆柱倾斜角, r 为圆柱半径.为了提高运算速度,假定倾斜角 α 始终为 0,则上述公式可简化为如下形式

$$\begin{cases} u_{ri} = \arcsin((u_i - u_0)/r) + u_0 \\ v_{ri} = v_i \end{cases} \quad (7)$$

柱面模型展开前后的人头部图象如图 2 所示.

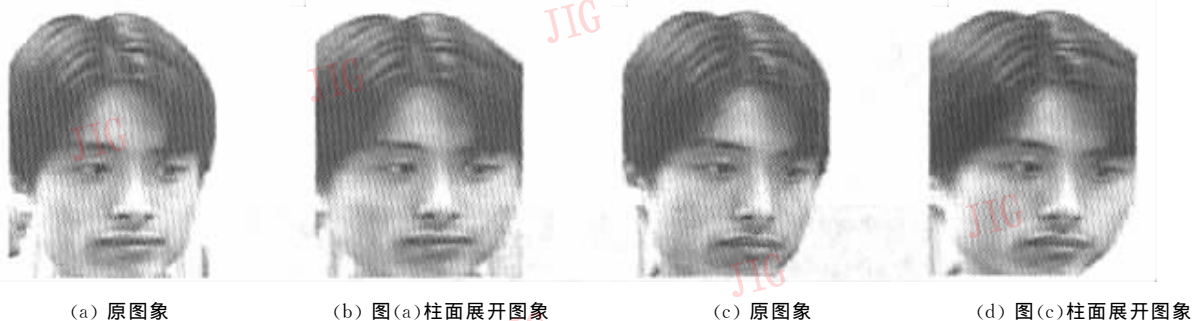


图 2 人头部圆柱模型展开图

实验结果表明,采用圆柱模型拟合人头部后,转身时,头部特征的跟踪率有明显提高.图 3 为 10 组实验数据的平均值,其纵轴为正确跟踪的特征数量,横轴为图象序列中的帧号,人转身运动的最大转角为 90°.

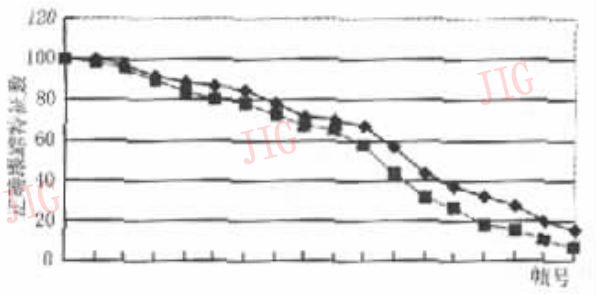


图 3 采用头部圆柱模型后转身时特征跟踪率的提高
(上面一组数据为使用圆柱模型后的结果)

1.3 2D 到 3D 人头部运动估计

假定人头部运动为刚体运动,当转动角度不大时,其表面特征点的运动方程如下

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & -\phi & \varphi \\ \phi & 1 & \theta \\ -\varphi & \theta & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (8)$$

其中, x, y, z, x', y', z' 分别为运动前后空间点在摄像机坐标系中的坐标, θ, φ, ψ 为特征点顺次绕 X, Y, Z 轴逆时针方向旋转的转角, t_x, t_y, t_z 为平移分量.根据式(5)正射投影模型,可得到图象中投影点的如下运动方程

$$\begin{cases} u' = u - \psi y + \varphi z + t_u \\ v' = v - \phi x + \theta z + t_v \end{cases} \quad (9)$$

其中, u, v, u', v' 分别为运动前后投影点的坐标.当人头部的俯仰和左右倾斜运动量不大时,可以忽略上式中的 θ, ϕ ,则该运动方程可简化为如下形式

$$\begin{cases} u' = u + \varphi z + t_u \\ v' = v + t_v \end{cases} \quad (10)$$

如果选择 n 个特征点,则可以得到 $2n$ 个运动方程.若利用最小二乘法来求解其中 3 个运动参数 φ, t_u, t_v ,即得到如下方程组

$$\begin{cases} \varphi \sum_i z_i^2 + t_u \sum_i z_i - \sum_i (u'_i - u_i) z_i = 0 \\ \varphi \sum_i z_i + nt_u - \sum_i (u'_i - u_i) = 0 \\ nt_v - \sum_i (v'_i - v_i) = 0 \end{cases} \quad (11)$$

其中, z_i 为特征点的深度值. 根据式(5)正射投影模型和式(7)头部圆柱简化模型, 可得如下该深度值求解公式

$$z_i = Z_0 - \sqrt{r^2 - (u_i - u_0)^2} \quad (12)$$

其中, Z_0 为人头部中心到摄像机的距离. 由于假定空间点深度值变化范围不大, 因此该值可设为一常数. 当将该式代入式(11), 即可解得 3 个运动参数. 这里需注意, 式中 u, v, u', v' 均为未施加柱面展开时的坐标.

2 基于头部几何特征的跟踪校正

上述块特征方法是通过跟踪窗口内批量块特征运动矢量的计算来实现目标跟踪的. 由于它仅利用

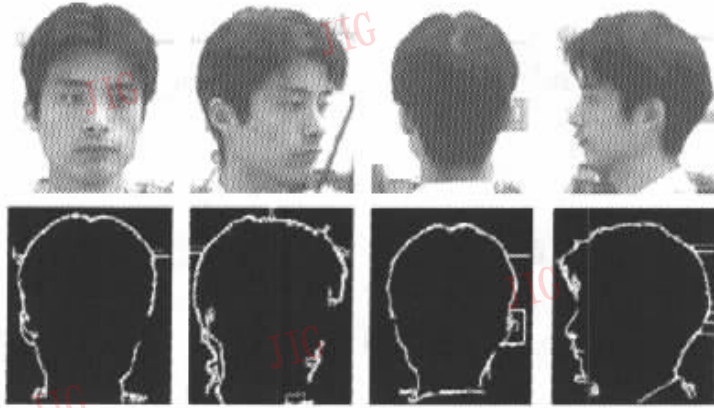


图4 人转身运动时, 头部的轮廓几何特征具有较强的相似性

跟踪窗口中头部轮廓偏移量的快速计算是该校正方法的前提, 为此, 本文引入迭代最近点匹配算法(ICP)来实现这一过程. 对于模型轮廓上的任一边缘点, 可在目标图象一定范围内搜索与它距离最近的边缘点, 并计算这两点之间的偏移, 所有这些偏移量的平均值定义为整体轮廓的偏移. 该过程可迭代运行, 以获取精确解, 其算法公式描述如下

$$[d_x, d_y] = \left[\frac{1}{n} \sum_i d_{x,i}, \frac{1}{n} \sum_i d_{y,i} \right] \quad (13)$$

其中, $d_{x,i}, d_{y,i}$ 为模型轮廓点到目标图象中最近边缘点的偏移, n 为模型轮廓点像素的个数.

该迭代最近点匹配算法不仅能够准确计算出两幅图象之间头部轮廓的偏移量, 而且对噪声、比例缩放、轮廓形变等具有很好的抵制能力. 该算法时间复杂度为 $O(nw^2k)$, 其中 n 为基准图象中轮廓边缘点的个数, w 为搜索窗口宽度, k 为迭代次数. 由于是在块特征跟踪的基础上进行校正, 搜索窗口宽度可

了图象低层信息, 因此在长时间跟踪时, 容易因误差累积而发生目标丢失的情况. 另外, 在跟踪过程中(尤其在人转身运动时)块特征需要不断更新, 但累积误差会造成更新的某些特征是非头部特征, 这也会导致目标逐渐丢失.

假设人头部在运动过程中为一刚体, 由于在不发生目标遮挡的情况下, 人平移或转身运动时, 头部轮廓特征十分显著(图4为人头部几种典型情况下的轮廓特征), 因此, 可以根据当前跟踪窗口中头部轮廓位置的偏移来对块特征跟踪的结果进行校正, 从而不仅解决了因误差累积等原因而造成的目标丢失问题, 而且实现了长序列图象中人平移或转身运动的准确跟踪.

以很小(一般为 10pixels), 因此其运行速度很快, 完全能满足实时校正的要求. 其中, 跟踪校正过程中的头部模型轮廓的提取可有多种方法, 如直接使用初始帧头部轮廓、提取多帧轮廓求中值、或提取多个模型轮廓等.

本文提出的基于几何特征的跟踪校正方法的另一好处是, 可以根据模型轮廓与当前帧边缘图象的匹配率来确定跟踪结果的置信值, 以实现错误跟踪情况的检测与处理.

3 实验结果

本文提出的人头部实时跟踪方法已在 P II 350 微机实现. 当跟踪窗口大小为 120×180 pixels, 块特征数为 80 时, 跟踪速度达到 30 帧/s(其中块特征跟踪处理时间约为 25ms, 跟踪校正时间约为 5ms), 相邻两帧头部最大位移为 30pixels. 实验结果表明,

本文算法不仅可以准确实时地跟踪人的平移运动,而且可以准确实时跟踪人的转身运动,同时在长时间跟踪时,不会因为误差累积等原因而发生目标丢

失的现象.另外,本文算法在光线较暗和光照条件变化较大等情况下也能实现可靠跟踪.图 5 为一次实时跟踪实验的几幅中间结果.



图 5 人平移及转身时头部运动的跟踪

4 讨论

本文主要讨论了一种使用黑白摄像机来实现人平移或转身时头部运动的实时跟踪方法,并取得了满意的实验结果.本文方法还可以跟踪头部的左右大幅度摇摆运动,但跟踪速度较慢,其主要原因是,在进行头部圆柱模型计算和跟踪校正计算之前,需要通过轮廓匹配来确定头部姿态,并以此确定运动跟踪参数,因这一计算过程耗时较大,故系统运行速度尚不能达到实时.解决这一问题的工作正在进行之中.

参考文献

- 1 Antonio C, Brendan F, Huang T S. Detection and tracking of faces and facial features[A]. In:Proc. IEEE on ICIP[C], Kobe, Japan, 1999:657~661.
- 2 Antonio C, Ricardo L, Huang T S. 3D model-based head tracking[J]. In:proc. of SPIE on VCIP[C], San Jose, CA, 1997, 3024:426~434.
- 3 Marco Cascia, John Isidoro, Stan Sclaroff. Head tracking via robust registration in texture map images[A]. In:Proc. IEEE on CVPR[C]. Santa Barbara,CA,1998:508~514.
- 4 Jie Yang, Alex Waibel. Tracking human faces in real-time[R]. CMU CS Technical Report, CMU-CS-95-210, 1995.
- 5 Nuria Oliver, Alex Pentland, Francois Berard. LAFTER:Lips and face real time tracker with facial expression recognition[A]. In:Proc. IEEE on CVPR[C]. San Juan, Puerto Rico,1997:123~129.
- 6 Cordea M D, Petriu E M, Georganas N D *et al.* Real-time 2^{1/2}D head pose recovery for model-based video-coding[A]. In:Proc.

IEEE on Instrumentation and Measurement Technology [C]. Baltimore, MD, 2000:601~606.

- 7 Crowley J L, Berard F. Multi-modal tracking of faces for video communications [A]. In:Proc. IEEE on CVPR[C]. San Juan, Puerto Rico,1997:640~645.
- 8 Jianbo Shi, Carlo Tomasi. Good features to track[A]. In:Proc. IEEE on CVPR[C]. Seattle, WA, 1994:593~600.
- 9 Tiziano Tommasini, Andrea Fusiello, Emanuele Trucco *et al.* Making good features track better[A]. In:Proc. IEEE on CVPR [C]. Santa Barbara, CA,1998:178~183.

徐一华 1978年生,北京理工大学硕士研究生.研究方向为计算机视觉、计算机图形学、智能系统、并行计算.

朱玉文 1950年生,北京理工大学计算机科学与工程系副教授.主要研究方向为模式识别以及图象处理.

贾云得 1962年生,工学博士,现任北京理工大学计算机科学与工程系教授,博士生导师.主要研究方向为计算机视觉、图象分析与理解、人工生命和智能系统.