

基于运动轨迹分析的头部行为识别

黄 腾, 阮宗才

(东南大学学习科学研究中心, 南京 210096)

摘要: 提出一种基于人脸运动轨迹分析的头部行为识别方法。对人脸进行检测和全自动跟踪, 通过计算人脸运动轨迹参数并进行参数数据统计和分析, 实现人体头部行为识别。针对抬头、低头时出现的跟踪框长短轴交换问题, 通过引入基于跟踪框长短轴比值的约束条件, 对椭圆跟踪盒主轴与纵轴夹角的计算方法进行改进。实验结果证明, 该方法的正确识别率较高。

关键词: 跟踪盒; 跟踪框; 质心; 主轴; 行为识别; 运动轨迹

Head Behavior Recognition Based on Analysis of Trajectory

HUANG Teng, RUAN Zong-cai

(Research Center for Learning Science, Southeast University, Nanjing 210096, China)

【Abstract】 A method of head behavior recognition based on the analysis of trajectory is presented. It detects and tracks the face automatically, and calculates parameters of the trajectory of the face. It can achieve the behavior recognition of the head by parameters' data statistics and analysis. For the problem of the exchange of the long axis and the short axis of the tracking window at the case of up head and down head, it proposes an improved algorithm of the angle between the major axis of the tracking box and the vertical axis. Experimental result shows that the problem of the exchange of the long axis and the short axis have been successfully resolved, and the algorithm can recognize the behavior of the head correctly.

【Key words】 tracking box; tracking window; centroid; major axis; behavior recognition; trajectory

DOI: 10.3969/j.issn.1000-3428.2012.17.059

1 概述

人的头部运动主要分为头部偏转、摇头、抬头和低头等行为。人的头部行为反映人的心理活动。如果能对人的头部进行行为跟踪, 那么不仅能应用于研究被试者的心理、情绪和态度^[1], 还能应用于医学治疗、人机接口等与机器视觉有关的领域。人脸跟踪系统利用计算机技术从包含人脸的图像序列中检测、分类和跟踪人脸, 并对其行为进行理解和描述。行为理解问题主要包括行为建模、行为识别和高层行为分析。行为建模的实时性和有效性最终决定行为理解的结果^[2]。

Gary R. Bradski 等使用 Camshift 对人脸进行跟踪, 并选取跟踪框的质心坐标、跟踪框长轴和短轴长度、跟踪盒主轴与纵轴的夹角为运动参数进行计算, 该算法可对人脸进行跟踪, 但未能对低头时出现的因跟踪框长短轴交换问题引起的头部偏转角度变化进行有效处理^[3], 当利用参数统计分析进行头部行为识别时, 该参数无效, 对抬头、低头等头部行为无法进行识别。针对该问题, 本文引入一种基于跟踪框长短轴比值的约束条件, 改进椭圆跟踪盒主轴与纵轴夹角的算法。此外, Gary R. Bradski 等仅提取出人脸的运动跟踪参数, 本文在此基础上, 采用跟踪框、跟踪盒运动参数数据进行基于人脸的运动轨迹分析方法的行

为建模, 利用该方法对头部行为如摇头、摆头、抬头等进行识别, 为心理分析提供基础。

2 运动目标检测

本文采用背景建模检测人脸。背景建模是运动物体检测中常用的一种方法, 即逐帧读取视频, 将前景图提取出来进行显示^[4]。把当前帧和前一帧各自的单通道灰度图进行差分并二值化后得到前景图:

$$D_k = |f_k(x, y) - B_{k-1}(x, y)| \quad (1)$$

$$F_k(x, y) = \begin{cases} 1 & D_k \geq T \\ 0 & D_k < T \end{cases} \quad (2)$$

即把前一帧设置为单通道图 $B_{k-1}(x, y)$, 当前帧转换成单通道帧图 $f_k(x, y)$, 当前帧与前一帧的灰度图进行差分并阈值化后得到前景图。背景建模提取的前景图如图 1 所示。

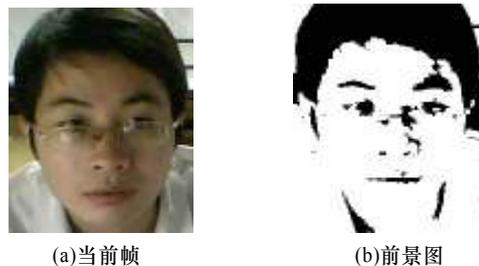


图 1 背景建模提取的前景图

作者简介: 黄 腾(1985—), 男, 硕士研究生, 主研方向: 计算机视觉, 目标跟踪; 阮宗才, 副教授、博士

收稿日期: 2011-11-17 **修回日期:** 2012-01-10 **E-mail:** huangteng08@126.com

3 全自动运动目标跟踪

在前景图的基础上, 自动提取人脸轮廓并设置初始跟踪框, 采用基于颜色特征和基于运动模型相结合的人脸跟踪算法, 算法描述如图 2 所示。

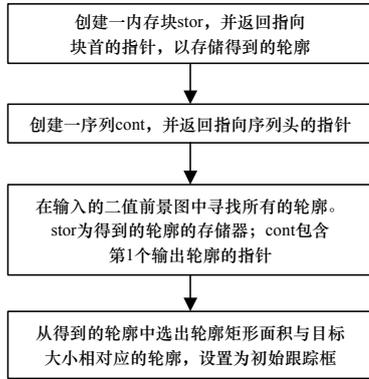


图 2 初始跟踪框提取流程

3.1 Camshift 算法

Camshift 算法^[3]是基于颜色特征的运动物体跟踪算法, 它以运动物体的颜色概率分布图为作用目标对其进行跟踪。Camshift 算法首先把 RGB 颜色空间转换成 HSV 颜色空间并建立运动物体一维 H 直方图。在操作过程中, 建立的一维 H 直方图作为一个查找表, 把每一帧的跟踪框区域转换成相应的颜色概率分布图, 然后使用以颜色概率分布图为目的的 Camshift 对运动物体进行跟踪。Camshift 算法步骤如下:

- (1)选择初始跟踪框的尺寸和位置。
- (2)把帧图转换到 HSV 颜色空间, 建立跟踪框中目标的一维 H 直方图。
- (3)根据一维 H 直方图, 建立目标的颜色概率分布图。
- (4)根据建立的颜色概率分布图计算跟踪框的质心, 并把跟踪框的中心移至质心。
- (5)重复步骤(4), 直至汇聚(或者是质心移动的距离小于某个预先设定的阈值), 记下 M_{00} 。
- (6)设置跟踪窗口的大小为 M_{00} 的一个函数。
- (7)重复步骤(4)~步骤(6), 直至汇聚。

对于离散的二维概率分布图, 跟踪窗质心的计算如下:

若:

$$M_{00} = \sum_x \sum_y I(x, y) \tag{3}$$

$$M_{10} = \sum_x \sum_y xI(x, y) \tag{4}$$

$$M_{01} = \sum_x \sum_y yI(x, y) \tag{5}$$

则跟踪窗的质心坐标为:

$$x_c = \frac{M_{10}}{M_{00}}; y_c = \frac{M_{01}}{M_{00}} \tag{6}$$

其中, $I(x, y)$ 是位置 (x, y) 处的像素值; x 和 y 属于跟踪窗口。

3.2 卡尔曼滤波

由于 Camshift 算法是基于颜色特征的跟踪算法, 因

此对光照比较敏感, 同时, 对于快速运动的物体也会出现跟丢目标的情况^[5]。采用卡尔曼预估器, 预测物体的下一运动状态, 可有效解决上述问题。

卡尔曼滤波器由一系列递归数学公式描述, 它们提供了一种高效可计算的方法来估计过程的状态, 并使估计均方误差最小^[6]。它可以估计信号的过去、当前甚至将来的状态, 即使不知道模型的确切性质。

离散时间过程由以下离散随机差分方程描述:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \tag{7}$$

其中, $x \in \mathbf{R}^n$ 为离散时间过程的状态变量。定义观测变量 $z \in \mathbf{R}^m$, 得到观测方程:

$$z_k = Hx_k + v_k \tag{8}$$

随机信号 w_k 和 v_k 分别表示过程激励噪声和观测噪声, 假设它们相互独立且服从正态分布:

$$p(w) \sim N(0, Q) \tag{9}$$

$$p(v) \sim N(0, R) \tag{10}$$

这里假设过程激励噪声协方差矩阵 Q 、观测噪声协方差矩阵 R 、增益矩阵 A 、 B 和 H 都为常数。

卡尔曼滤波器分为两部分: 时间更新过程和测量更新过程^[7]。时间更新过程及时向前推算当前状态变量和误差协方差估计的值, 并投射到测量更新过程, 其方程描述为:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \tag{11}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{12}$$

测量更新过程校正先验估计以获得状态的后验估计, 并作为下一次的先验估计, 其方程描述如下:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{13}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \tag{14}$$

$$P_k = (I - K_k H)P_k^- \tag{15}$$

其中, \hat{x}_k^- 为当前状态变量的先验估计值; P_k^- 为先验估计误差的协方差; K_k 为残余增益。

本文利用卡尔曼滤波器预估跟踪框质心的坐标。各矩阵的选取为: 状态向量为质心的横、纵坐标的 2×1 维矩阵 $x_k = \begin{pmatrix} x \\ y \end{pmatrix}$, 系数矩阵 $A[] = \{1, 1, 0, 1\}$; 控制向量 $u_k = 0$; 过程噪声矩阵 $Q[] = \{1e-5, 0, 0, 0\}$; 测量向量 $z_k = (x)$, 只对质心横坐标测量, 测量矩阵 $H[] = \{1, 0, 0, 0\}$; 测量噪声矩阵 $R[] = \{1e-1, 0, 0, 0\}$ 。

4 行为参数统计和行为识别

对运动物体实现跟踪后, 接下来就是对其行为进行分析和描述^[8]。本文选择跟踪框的质心坐标、长短轴的长度及其比值、椭圆跟踪盒主轴相对于纵轴的偏移角度作为参数, 通过对这些参数数据进行统计和分析实现对人体头部行为的识别。

各个参数的算法描述如下(质心的方程见式(6)):

若:

$$M_{20} = \sum_x \sum_y x^2 I(x, y) \tag{16}$$

$$M_{02} = \sum_x \sum_y y^2 I(x, y) \tag{17}$$

$$M_{11} = \sum_x \sum_y xyI(x,y) \tag{18}$$

则目标主轴偏转的角度为：

$$\theta = \frac{\arctan \left(\frac{2 \left(\frac{M_{11}}{M_{00}} - x_c y_c \right)}{\left(\frac{M_{20}}{M_{00}} - x_c^2 \right) - \left(\frac{M_{02}}{M_{00}} - y_c^2 \right)} \right)}{2} \tag{19}$$

其中，(x,y)属于跟踪窗口；I(x,y)为(x,y)处的像素灰度值。θ在图中位置如图3所示，图3中，向右偏为负，向左为正。跟踪框、跟踪盒示意图如图4所示。

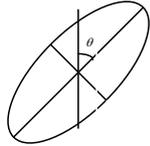


图3 主轴相对于纵轴偏转角度示意图

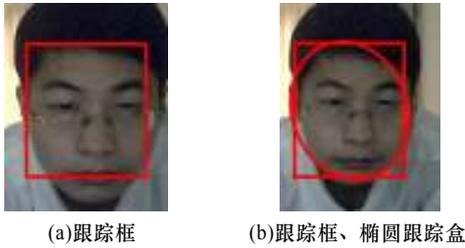


图4 跟踪框、跟踪盒示意图

若：

$$a = \frac{M_{20}}{M_{00}} - x_c^2 \tag{20}$$

$$b = 2 \left(\frac{M_{11}}{M_{00}} - x_c y_c \right) \tag{21}$$

$$c = \frac{M_{02}}{M_{00}} - y_c^2 \tag{22}$$

则跟踪窗口的长轴l和短轴w分别为：

$$l = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \tag{23}$$

$$w = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}} \tag{24}$$

那么，长轴和短轴长度的比值为l/w。

本文在文献[2]方法的基础上，改进当跟踪框长短轴发生交换时头部偏转角度的计算方法。低头时跟踪框长轴和短轴会发生交换，即长轴与短轴的比值小于1，短轴变成主轴，当原有算法得到的偏转角度小于零度时(即向右偏转)，偏转角度加上90°得到新的角度，当原有算法得到的偏转角度大于等于0°时(即向左偏转)，偏转角度减去90°得到新的角度，从而补偿了主轴的转换，得到比较理想的实验结果。即当满足约束条件l/w < 1时，补偿后主轴偏转的角度为：

$$\theta' = \begin{cases} \theta + 90 & \theta < 0 \\ \theta - 90 & \theta \geq 0 \end{cases} \tag{25}$$

5 实验结果与分析

本文实验使用 Intel 公司的开源项目——OpenCV 实现。

实验首先对人脸进行实时跟踪，然后统计每一帧中跟踪框的质心坐标、长短轴长度、跟踪盒主轴与纵轴的夹角等参数数据及其分布，最后通过各参数曲线分布图实现头部行为识别。图5为提取出的初始跟踪区域的颜色直方图。摆头、低头、摇头状态如图6所示。图6瞬时状态参数如表1所示。



图5 初始跟踪区域物体颜色直方图

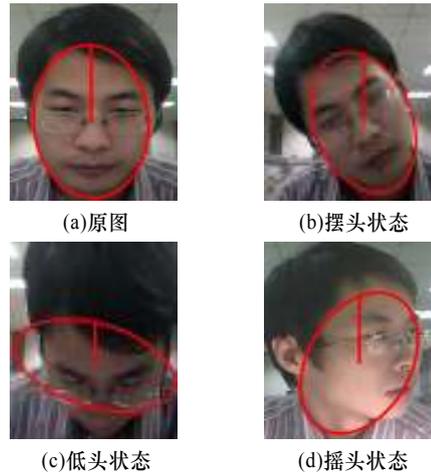


图6 各种姿态下跟踪图

表1 图6中各瞬时状态参数数据

图像	短轴	长轴	质心坐标	夹角/(°)
图6(a)	161	251	(343, 202)	2.819 572
图6(b)	138	199	(351, 187)	24.856 895
图6(c)	252	116	(336, 152)	-12.520 475
图6(d)	140	221	(374, 188)	-22.940 239

表2~表4统计了摆头、低头和摇头状态下的各个参数数据分布，其中，帧(frames)为第几帧帧数，短轴(width)、长轴(height)分别为跟踪框的短轴w和长轴l，质心坐标(centroid)为跟踪框质心坐标，夹角(angle)即椭圆跟踪盒主轴与纵轴的夹角。

表2 摆头状态下各种行为的参数数据统计

帧	短轴	长轴	质心坐标	夹角/(°)
33	170	273	(365,239)	-89.499 290
34	172	270	(378,236)	-81.091 011
35	162	251	(399,226)	-65.671 455
36	158	239	(401,217)	-63.302 685
37	152	223	(401,199)	-63.958 717
38	143	212	(399,200)	-62.930 698
39	143	216	(394,203)	-66.577 950
40	147	228	(391,199)	-69.393 097
41	154	235	(391,207)	-68.743 958
42	156	234	(392,233)	-64.226 860
43	169	242	(377,246)	-73.693 810
44	178	246	(254,246)	88.823 349
45	173	240	(343,243)	79.712 402
46	160	217	(324,234)	62.891 632

表3 低头状态下各种行为的参数数据统计

帧	短轴	长轴	质心坐标	夹角/(°)
26	181	216	(359,209)	72.366 341
27	156	178	(359,196)	55.830 742
28	175	125	(347,166)	29.677 416
29	207	109	(354,138)	15.029 250
30	229	99	(357,90)	11.314 209
31	226	95	(359,55)	7.346 985
32	115	128	(452,64)	-48.089 550
33	78	203	(451,101)	-72.442 673
34	115	195	(447,147)	88.896 561
35	142	205	(349,235)	-81.674 034
36	142	178	(349,247)	-71.532 722
37	166	114	(336,125)	31.065 578
38	209	130	(343,172)	10.361 908
39	219	138	(347,184)	9.748 791
40	208	149	(346,204)	12.139 671

表4 摇头状态下各种行为的参数数据统计

帧	短轴	长轴	质心坐标	夹角/(°)
26	181	216	(359,209)	72.366 341
27	156	178	(359,196)	55.830 742
28	175	125	(347,166)	29.677 416
29	207	109	(354,138)	15.029 250
30	229	99	(357,90)	11.314 209
31	226	95	(359,55)	7.346 985
32	115	128	(452,64)	-48.089 550
33	78	203	(451,101)	-72.442 673
34	115	195	(447,147)	88.896 561
35	142	205	(349,235)	-81.674 034
36	142	178	(349,247)	-71.532 722
37	166	114	(336,125)	31.065 578
38	209	130	(343,172)	10.361 908
39	219	138	(347,184)	9.748 791
40	208	149	(346,204)	12.139 671

图7~图13显示了各种头部行为状态下各个参数数据随帧数的分布曲线。

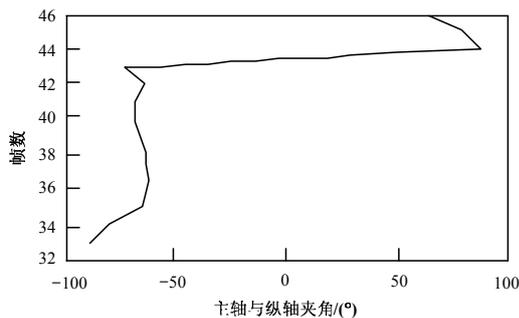


图7 摇头状态下夹角分布曲线

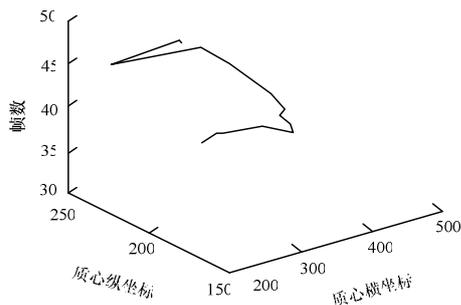


图8 摇头状态下质心分布曲线

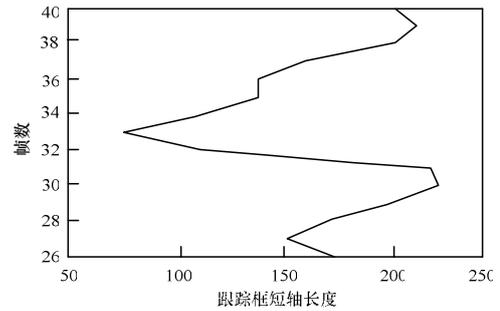


图9 低头状态下跟踪框短轴分布曲线

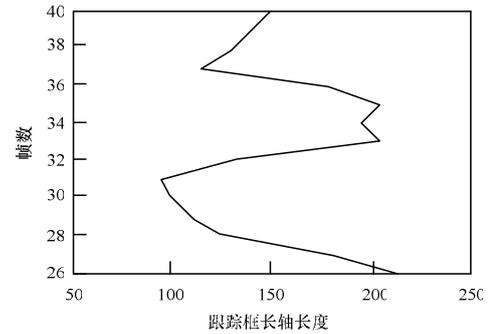


图10 低头状态下跟踪框长轴分布曲线

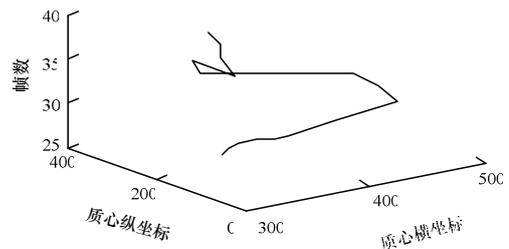


图11 低头状态下质心分布曲线

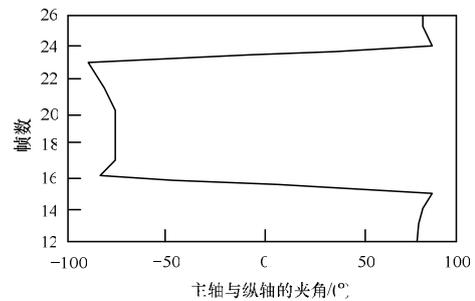


图12 摇头状态下夹角分布曲线

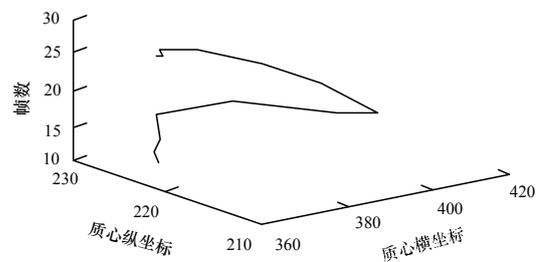


图13 摇头状态下质心分布曲线

限于篇幅,本文只给出3种头部行为状态参数及曲线分布。最后结果证明,本文基于人脸运动轨迹分析的头行为识别方法,即对人体头部的各种简单行为及其之间的
(下转第231页)