

文章编号: 1001 - 9081(2012) S1 - 0128 - 06

基于 OpenCV 的摄像头动态手势轨迹识别及其应用

江 超, 艾娇燕*

(广西大学 电气工程学院, 南宁 530004)

(* 通信作者电子邮箱 a520jc@163.com)

摘 要: 利用 OpenCV 计算机视觉库在 vs2008 平台上设计了一个基于实时摄像头的集动态手势检测、动态手势跟踪、动态手势轨迹识别的应用。首先, 该应用基于静止的背景更新, 利用背景差分检测运动手势, 再结合颜色直方图的粒子滤波进行动态手势跟踪, 最后利用隐马尔可夫模型(HMM) 进行运动轨迹识别。在运动检测部分结合了背景差分图与通过颜色直方图获得的反投影图, 达到比较满意的实时运动检测效果; 在运动手势跟踪部分, 改进的颜色直方图的粒子跟踪能够在经过类肤色人脸的干扰后迅速地找回运动手势, 基本达到了跟踪的要求, 但是同时对于 HMM 识别轨迹时需要的运动轨迹序列采集造成了影响; 在识别轨迹部分, HMM 的训练达到了识别的要求, 但是识别的效果主要取决于实时运动轨迹序列的采集工作与采集方法的优化。

关键词: OpenCV; 摄像头; 粒子跟踪; HMM; 轨迹识别

中图分类号: TP391.41 **文献标志码:** A

Dynamic gesture trajectory recognition and its application design for camera based on OpenCV

JIANG Chao, AI Jiao-yan*

(School of Electrical Engineering, Guangxi University, Nanning Guangxi 530004, China)

Abstract: Using OpenCV computer vision library, the authors designed an application in vs2008 platform, which was based on real-time camera and made up of dynamic gesture detection, dynamic gesture tracking, dynamic gesture trajectory recognition. Firstly, the application based on static background update, used the background subtraction to detect motion gestures; secondly, it used particle filter combined with the color histogram for dynamic gestures tracking; and then, it used the Hidden Markov Model (HMM) for trajectory recognition. In part of motion detection, combined with the projection map from color histogram and background subtraction, it can get satisfactory effect of real-time motion detection; in part of motion gesture tracking, the particle tracking with improved color histogram can quickly get back the motion gestures when going through interference in color of skin-like face, which basically meet the tracking requirements but influences the collection of HMM trajectory sequence; in part of trajectory recognition, the training of HMM meets the requirements of recognition, but the effect of recognition depends on the collection works of real-time trajectory sequence and the the optimization of collection method.

Key words: OpenCV; camera; particle tracking; Hidden Markov Model (HMM); trajectory recognition

0 引言

动态手势识别的一个重要用途是人机交互, 而且对智能控制和视频监控等领域都有十分重大的意义。基于实时摄像头的手势识别更增加了人机交互的可能性, Intel 也为计算机视觉领域的应用开发提供了开源的计算机视觉库。

动态手势识别包括手指运动和手指不运动的两种动态手势, 本文的工作是在个人电脑上实现简单的人机交互, 只对手的运动轨迹进行识别。在简单的人机交互过程中, 背景没有运动, 因而选择静止背景的运动检测^[1]来检测运动手势; 在手势跟踪方面, 随着计算机性能的提高和成本的下降, 选择具有非线性的计算量偏大的粒子滤波跟踪^[2]算法; 在识别方面采用在语音识别上取得巨大成功的隐马尔可夫(Hidden Markov Model, HMM) 识别算法^[3]。

1 OpenCV 简介

OpenCV 是 Intel 开源的计算机视觉库(open source computer vision library) 的简称, 它是由一些 C 函数和少量的 C++ 函数构成, 实现了在图像处理以及计算机视觉方面的许多通用算法, 主要对图像进行如运动检测、跟踪、分割、识别以及 3D 重建等高级处理。它有 300 多个 C 函数的跨平台的中、高层 API, 它不依赖外部库, 但可以使用外部库。由于它是开放源码且代码简洁高效, 其中大多数函数都经过汇编最优化, 所以能充分利用英特尔系列的芯片的设计体系, 因而对于英特尔系列芯片, OpenCV 的代码执行效率是非常高的。OpenCV 具有强大的图像和矩阵运算能力和这方面的用户接口, 同时支持 Windows 和 Linux 系统, 本文程序运行环境为: Windows + VS2008 + OpenCV2.1 + DirectShow9.0 + GSL1.8。

收稿日期: 2011-11-08; 修回日期: 2011-12-28。

作者简介: 江超(1985 -) 男, 湖北武汉人, 硕士研究生, 主要研究方向: 模式识别与智能控制、计算机视觉; 艾娇燕(1970 -) 女, 湖南株洲人, 副教授, 博士, 主要研究方向: 模式识别、计算机视觉、智能检测与控制。

2 基于颜色直方图与背景差分的运动检测

背景差分是将当前帧图像和背景图像相减得到前景图像,但在实际摄像头采集图像过程中背景会随光照以及人脸和身体的轻微抖动而不断发生微小变化,如果不对背景图像进行重新初始化,错误信息会不断地累积,从而不能有效地检测出运动目标。而由于程序在启动摄像头过程中需要一定时间,在这段时间里,摄像头差不多可以采集 20 帧的图像,且人在这段时间基本是等待摄像头开启,因此对这 20 帧进行循环背景更新^[4]如式(1)并对每帧运动检测后得到除运动手势区域外的像素也做背景更新。

$$\mu_{t+1} = \begin{cases} \mu, & I_f(x, y) = 1 \\ \alpha\mu_t + (1 - \alpha)I_t, & \text{其他} \end{cases} \quad (1)$$

其中: μ_t 为背景图像的相应的像素点, μ_{t+1} 为更新后的背景图像像素点; I_t 为当前帧图像的像素点, I_f 是当前帧图像像素点的掩码值,即是否做背景更新; $\alpha \in [0, 1]$ 为背景模型更新速度,一般取 0.8 到 1 之间,本文取 0.8。OpenCV 中可用 `cvAddWeighted` 函数实现上述背景更新;在摄像头启动获得背景图像后,再对每帧的当前图像与背景图像差分得到 BGR 的运动前景图像,在 OpenCV 中可以用 `cvAbsDiff` 函数实现实现背景差分,此时的差分图像为 3 通道的 BGR 图像,要得到运动二值图需转换到单通道的灰度图,用 OpenCV 的 `cvCvtColor` 函数可以转换到灰度图。在得到运动灰度图后本可以设定阈值就可以得到运动二值图,在 OpenCV 中用 `cvThreshold` 实现二值图的转换,但这样得到的运动手势图效果不好,本文尝试通过颜色直方图获得当前帧图像的反投影图再经过二值化也得到一个二值图。由于 HSV 空间对光的敏感程度较小,因而颜色直方图采用 HSV 空间。方法如下:

由于黄种人的肤色偏黄色,纯黄色是红色和绿按 1:1 搭配而得,所以在 RGB 图像中,感兴趣的肤色区域的像素其红色一定大于绿色即 $r > g$,由于本实验采用 30 万的摄像头,采集肤色难免有失真,在多次调试发现该摄像头,在室内环境内采集肤色的像素其红色与绿色的比例大约在 1.0 和 2.0 之间。所以同时满足运动且 r/g 比例在 1.0 ~ 2.0 的条件的像素点最有可能是运动的肤色。而是肤色的像素点的亮度也比较大,在满足上述条件的像素点中,对其在 HSV 空间中对亮度 v 值重新设置为比较小的亮度值(减小静止类肤色的干扰)。于是可以在采集 HSV 颜色直方图的过程中,对满足上述肤色条件的像素对应的亮度 v 都做此设定,而在实际检测每帧图像时都对满足运动肤色条件的像素的亮度 v 也做同样的处理,这样通过人为修改过的直方图得到每帧图像的反投影图像,再经过二值化,就可以通过人为的标记检测出运动的手势二值图像(受静止类肤色的干扰影响小)。对 HSV 空间的 v 重新设定如式(2):

$$\begin{cases} v_temp = \begin{cases} r/g, & g \neq 0 \\ 4, & g = 0 \end{cases} \\ v = \begin{cases} 10r/g, & v_temp \leq 2 \\ 255, & v_temp > 2 \end{cases} \end{cases} \quad (2)$$

其中 $r > g$ 。然后将得到的运动二值图和通过反投影得到二值图像相与,再进行一些图像形态学闭操作处理可以得到比较完

整的运动肤色手势二值图。在 OpenCV 中用 `cvAnd` 函数实现两个图像的相与操作,用 `cvMorphologyEx` 函数实现形态学操作。

3 颜色直方图与运动检测结合的粒子跟踪

3.1 粒子滤波原理及算法

粒子滤波就是初始时刻产生一些随机的粒子,这些粒子随时间传播,去寻找符合要求的最优解,假如在 $t-1$ 时刻每个粒子都有一个解,每个解与真实解之间有一个权重,代表与真实解的接近程度,随着时间增加,越接近真实解的粒子权重越来越大,而远离真实解的粒子权重变得越来越小,最后趋于 0,从而最终找到真实解的一种概率方法(如图 1)。

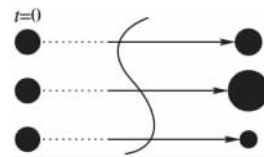


图1 粒子滤波的基本原理

粒子滤波的基本算法有贝叶斯重要采样 (Bayesian Importance Sampling, BIS), 序列重要采样 (Sequence Importance Sampling, SIS)^[5]。BIS 是说对于求解贝叶斯后验概率过程中,在系统真实的后验概率密度分布中采样比较困难,从而需要从一个已知的,易于采样的概率密度分布采样,这个概率分布称为建议分布。原理如式(3)^[6]所示:

$$E[g(x)] = \int g(x) p(x) dx = \int \frac{g(x)p(x)}{q(x)} q(x) dx \quad (3)$$

于是求解 $g(x)$ 的期望就从 $p(x)$ 分布转到了 $q(x)$ 分布了。SIS 的原理是要使粒子在 k 时刻的权值依赖于 $k-1$ 时刻的权值和数据,使粒子权值成为一种时间序列,从而能更好地估计粒子权值,迭代计算粒子的重要性加权如式(4)^[6]:

$$w_k^i(x_{0:k}) = w_{k-1}^i(x_{0:k}) \frac{p(z_k | x_k^i) p(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{k-1}^i, z_{1:k})} \quad (4)$$

其中 $w_{k-1}^i(x_{0:k})$ 是第 i 个粒子 $k-1$ 时刻的权重, $w_{k-1}^i(x_{0:k})$ 后面分式包含 $k-1$ 时刻的数据信息。选择建议分布为状态转移先验概率分布^[6]即 $q(x_k | x_{k-1}, z_k) = p(x_k | x_{k-1})$ 时,式(4)简化如式(5):

$$w_k^i = w_{k-1}^i * p(z_k | x_k^i) \quad (5)$$

即 k 时刻的粒子权值等于 $k-1$ 时刻粒子的权值乘以系统观测的似然概率^[6],由于序列重要采样会产生粒子退化现象^[6]即经过若干次迭代后,极少数粒子会拥有较大权值,而其他粒子则分配到较小权值,这样退化的粒子集无法有效表达后验概率,而且也会浪费时间在计算这些对后验概率贡献较小的粒子上,于是要对退化的粒子进行重采样^[7],重采样对退化较严重的粒子舍弃,同时将拥有较大权值的粒子复制给该退化的粒子,这样就保证了粒子的数量不会改变,而且使粒子集中趋向于真实的解(如图 2)。

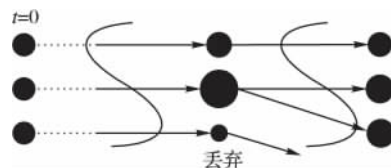


图2 粒子滤波重采样原理

3.2 基于改进直方图的粒子滤波算法

3.2.1 粒子直方图的算法

基于颜色直方图的粒子滤波就是将图像颜色特征的相似度作为粒子滤波要求解的后验概率,通过计算颜色直方图的巴氏系数 Bhattacharyya^[8] 如式(6) 来得到粒子的权重:

$$p[p, q] = \sum_{u=1}^m \sqrt{p(u)q(u)} \quad (6)$$

$$d = \sqrt{1 - p[p, q]} \quad (7)$$

巴氏距离^[9]就相当于两个向量夹角的余弦(如图3)。

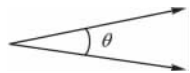


图3 两个向量的夹角

θ 越小相似度越高,其中 m 为直方图的颜色分组的 bin 格数, $p(u)$ 为参考手势颜色概率分布, $q(u)$ 为候选粒子所在区域颜色概率分布, d 为巴氏距离。

将巴氏距离作为每个粒子所在的图像区域的颜色直方图与采集的手势颜色直方图的相似度,来判断该粒子所在的区域是手势的程度,因而巴氏距离就作为每个粒子的权重。

由于 HSV 色彩空间对光线敏感程度小,粒子直方图广泛选用 HSV 空间,一般为了减小光照的影响,对直方图的 V 分量划分得比较粗,本文采用在运动检测时对 HSV 的 V 分量做第2章中相同的处理,这样在在计算直方图的相似度时对运动的肤色更有聚类性,也减小了非运动的肤色对粒子跟踪的干扰。

本文粒子直方图的基本算法流程是:

- 1) 首先初始化粒子数为 70,采集视频图像中手部区域的直方图作为参考手势直方图,并保存直方图数据。
- 2) 判断是否完成手势直方图的采集:如果是,则初始化粒子中心、粒子矩形区域和粒子权重;否则回到 1) 继续采集手势直方图。
- 3) 判断粒子是否完成初始化,如果完成粒子初始化,就进行粒子随机传播,然后在下一帧图像中计算每个粒子所在矩形区域内的直方图;否则等待继续完成初始化。
- 4) 在步骤 3) 中完成了每个粒子的直方图计算后,通过式(4)(5) 计算各个粒子直方图与参考手势直方图的巴氏距离,通过巴氏距离更新粒子权重,并将粒子权重归一化。
- 5) 对权重较小的粒子重采样。
- 6) 将粒子权重作为粒子中心出现的概率,计算所有粒子中心的期望,即可能跟踪的手势中心。
- 7) 回到 3) 进入下一帧图像循环。

图4 是粒子状态转移的传播模型^[10]。

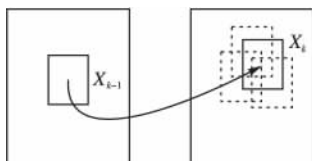


图4 粒子状态转移的传播模型

每个粒子在初始化时都被赋予相同的状态,然后每个粒子经过随机传播,随机地被赋予不同状态和位置信息。本文采用随机的传播模型如式(9):

$$X_{k+1} = A_1 * (X_k - X_0) + A_2 * (X_{k-1} - X_0) +$$

$$B_0 * random + X_0 \quad (9)$$

其中: X_{k+1}, X_k, X_{k-1} 分别为当前粒子下一帧图像将要传播到的粒子坐标、当前帧图像的粒子坐标、粒子初始化时的坐标; A_1, A_2, B_0 为传播系数; $random$ 为随机数,本文用 GNU 的数学函数库 `gsl` 的 `gsl_ran_gaussian` 随机函数产生随机数。上式以粒子的初始坐标为起点开始传播。

由于粒子传播是随机的,所以可以到达图像中的每一个坐标,而与整个图像相比,运动手势区域相对较小。由于运动检测中已经可以检测出运动手势,并计算可以获得一个先验中心坐标,为了减小计算量,于是本文将粒子传播的区域设定在以该中心坐标为中心的大小为整个图像的 $1/3$ 矩形区域内,这样既包含了运动手势区,也使所有的粒子都只会在比整幅图像相对较小区域内,因而可以通过减小粒子来减小计算量。这样经过重采样后权重较大的粒子也会比较多,然后再计算所有粒子中心相对于权重的期望中心,即所有粒子中心平均值如式(8) 从而获得期望的最优粒子,并显示该粒子。

$$\begin{cases} x_k^{p_center} = \sum_{i=1}^N x_k^i * w_k^i \\ y_k^{p_center} = \sum_{i=1}^N y_k^i * w_k^i \end{cases} \quad (8)$$

其中: x_k^i 为第 k 帧的第 i 个粒子的横坐标, y_k^i 为第 k 帧第 i 个粒子的纵坐标; w_k^i 为第 k 帧得到的粒子权重; N 为粒子总数; $x_k^{p_center}, y_k^{p_center}$ 得到的所有粒子中心的平均坐标。如图4中虚矩形框为单个粒子,实矩形框为最终得到的期望粒子。

4 HMM 轨迹识别

4.1 HMM 的原理与概念

隐马尔可夫模型^[11]是一种基于参数表示用来描述随机过程的统计特性的概率模型。它是一个双重随机过程,由两个部分组成:一个部分是用转移概率描述状态转移的马尔可夫链^[11],叫作隐含层;另一个部分是用观测概率描述状态与观察序列间的关系,叫作观测层。一个 HMM 由一个五元素的参数 $\lambda = (S, O, A, B, \pi)$ 来表示,简称为 $\lambda = (\pi, A, B)$ 。它们的含义如下:

S : 状态的有限集合 $S = \{s_1, s_2, \dots, s_N\}$, 其中 N 为隐含的状态维数。

O : 观察值的有限集合 $O = \{o_1, o_2, \dots, o_T\}$, 其中 T 为观察序列的时长。

A : 与时间无关的状态转移矩阵 $A = \{a_{ij} | i, j = 1, 2, \dots, N\}$ 即 $a_{ij} = P(q_t = S_j | q_{t-1} = S_i)$, 它满足 $\sum_{j=1}^N a_{ij} = 1$ 。

B : 描述观测值的概率分布的混淆矩阵 $B = \{b_j(k) | j = 1, 2, \dots, N; k = 1, 2, \dots, M\}$, 它满足 $\sum_{k=1}^M b_j(k) = 1$ 其中 M 为观察值的总数。

π : 用来描述观察序列 O 在 $t = 1$ 时刻的初始概率分布, $\pi = [\pi_1, \pi_2, \dots, \pi_N]$ 即 $\pi_i = P(q_1 = S_i) | i = 1, 2, \dots, N$ 。

要利用 HMM 进行手势识别,必须解决三个基本问题:

1) 在给定观察序列 $O = \{o_1, o_2, \dots, o_T\}$ 和模型 λ 的情况下,如何计算 $P(O | \lambda)$ 。

2) 在给定观察序列 $O = \{o_1, o_2, \dots, o_T\}$ 和模型 λ 的情况

下,如果选择一个状态序列 $S = \{s_1, s_2, \dots, s_T\}$ 来最好地解释观察序列 O 。这是 HMM 的识别隐形状态链的问题。

3) 如何调整模型参数 $\lambda = (\pi, A, B)$ 使得 $P(O|\lambda)$ 最大。

4.2 HMM 的参数重估算法与识别

本文主要涉及运动轨迹的识别问题,于是重点阐述第三个问题,即通过训练如何调整模型参数,来对特定的轨迹进行识别。对于 HMM 样本训练的问题,主要是通过参数重估公式来估计 HMM 的参数^[3]。

$$\gamma_t(i) = P(q_t = s_i | o, \lambda) =$$

$$\alpha_t(i) \beta_t(i) / \sum_{i=1}^N \alpha_t(i) \beta_t(i) \quad (10)$$

式(10)中 $\gamma_t(i)$ 表示 t 时刻位于隐藏状态 s_i 的概率变量,其中前向变量 $\alpha_t(i) = P(o_1 o_2 \dots o_T, q_t = s_i | \lambda)$, 后向变量 $\beta_t(i) = P(o_{t+1} o_{t+2} \dots o_T, q_t = s_i | \lambda)$, 且有 $\sum_{i=1}^N \gamma_t(i) = 1$ 。对于给定观察状态 O 和模型 λ , 定义在 t 时刻处于隐藏状态 s_i 在 $t+1$ 时刻处于隐藏状态 s_j 的概率变量为:

$$\begin{aligned} \xi_t(i, j) &= P(q_t = s_i, q_{t+1} = s_j | o, \lambda) = \\ &= \alpha_t(i) a_{ij} \beta_{t+1}(j) / P(o | \lambda) = \\ &= \alpha_t(i) a_{ij} \beta_{t+1}(j) / \sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} \beta_{t+1}(j) \end{aligned} \quad (11)$$

且有 $\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j)$ 。于是通过式(10)和(11)可以给出一组合理的重估公式:

$$\bar{\pi}_i = \text{在时刻}(t=1) \text{ 处于状态 } s_i \text{ 的概率} = \gamma_1(i) \quad (12)$$

$$\bar{a}_{ij} = \frac{\text{从状态 } s_i \text{ 过渡到状态 } s_j \text{ 的平均次数}}{\text{从状态 } s_i \text{ 向其他状态转移的平均次数}} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \quad (13)$$

$$\bar{b}_j(k) = \frac{\text{处于 } s_j \text{ 和出现观察值 } O_k \text{ 平均次数}}{\text{处于状态 } j \text{ 的平均次数}} = \frac{\sum_{t=1, o_t=O_k}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)} \quad (14)$$

其中 $\bar{\pi}, \bar{a}_{ij}, \bar{b}_j(k)$ 分布表示 HMM 经过重估的 π, A, B 三个参数,式(12)~(14)组成了著名的 Baum-Welch 算法^[12]。

4.3 对运动轨迹的 HMM 观察值编码

要利用 HMM 对手势中心的轨迹进行识别,就要必须对手势中心轨迹的点进行特征提取并对定义 HMM 的观察值。而对于一个由手势中心点组成的运动轨迹一般包含三个信息:轨迹的初始点、轨迹的运动方向、轨迹运动的距离。对于简单的运动轨迹,用轨迹的方向向量就可以识别一个轨迹,因此可将轨迹的运动方向向量作为 HMM 的观察值并进行方向划分。将运动轨迹的方向夹角量化成 8 等份,如图 5。

通过判断轨迹点移动的方向,就可以识别手势如向左、向上、向右、向下等的含义,如果对轨迹夹角加以细分,必要时通过判断轨迹移动的距离,可以识别更多的轨迹含义诸如圆形轨迹、三角形等。本文只对简单含义的轨迹进行识别(如图 6)。

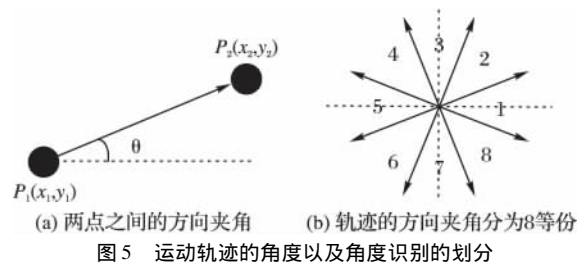


图 5 运动轨迹的角度以及角度识别的划分

本文通过 Baum-Welch^[3] 算法对四个方向的轨迹序列进行训练,将训练得到的四个方向的模型存为文本文件,然后在粒子跟踪算法中对得到的最优手势中心进行采集,计算连续两个手势中心的移动角度,然后按照图 5 所示加以相应的编码,保存在一个序列中,当采集完序列长度后,就将其作为四个 HMM 的输入序列计算相应的前向概率,得到最大概率的 HMM,就是最终得识别结果,并输出相应的控制信号,从而达到识别和控制的目的。

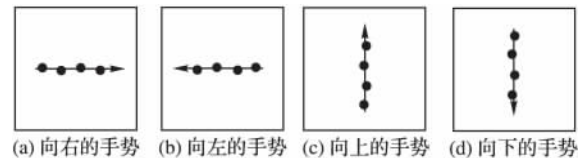


图 6 右、左、上、下的轨迹识别

由于摄像头采集具有实时性,因而何时开始采集序列,何时停止采集,是必须考虑的问题。本文经过多次实验调试且出于简单控制考虑,采取以下方法:本文设定了三个计时器,分别是开始准备采集序列计数器 timer0、当手势停留在图像中某个位置不动时的计数器 timer1、当手势消失在图像中时的计数器 timer2。手势序列开始采集的条件为:当手势停留在图像中心的一个长宽都为图像长宽 1/4 的矩形内, timer0 开始计数,当 timer0 等于 10 帧的时候,开始采集轨迹角度序列;手势序列停止并准备重新采集的条件为:1) 当手势开始采集且移动到图像外面,且序列采集还没完成时, timer2 开始计数,且当 timer2 等于 4 帧时,重新开始采集;2) 当手势开始采集且在采集过程中停在图像某处不动且序列采集还未完成时,并判断前后两帧手势中心移动的距离是否小于 3 个像素单位,如果是 timer1 开始计数,且当 timer1 = 3 帧时,重新开始采集。

5 实验结果与分析

5.1 基于背景差分的运动检测实验

从图 7 可以看出在连续的几帧中,有大量的运动的手势被检测出来,而人体微小的运动也被检测出来。

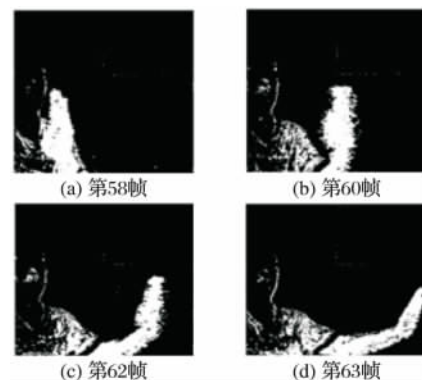


图 7 背景差分的运动检测

5.2 基于人为标定的 HSV 直方图的运动检测实验

从图 8 可以看出在连续的几帧中,经过人为标定的手势轨迹能有效地被检测出来,而没有经过编码的人脸部分,基本没有被检测出来。

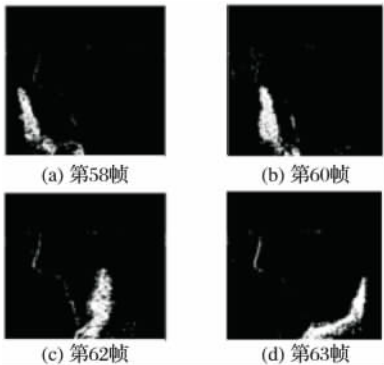


图 8 颜色直方图的运动检测

如图 9 在经过背景差分 and 人为标定直方图相结合的运动检测后,经过图像形态学处理能得到比较完整的运动手势二值图像。

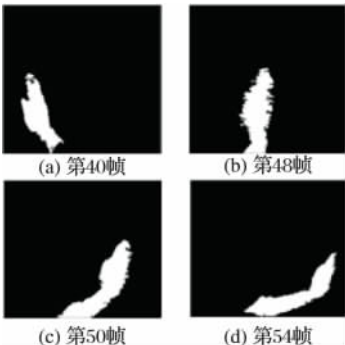


图 9 背景差分和颜色直方图相结合的运动检测

表 1 向右和向左的 HMM 部分状态转移的概率数据以及初始状态概率

模型状态	状态 1(右)	状态 2(右)	状态 1(左)	状态 2(左)	π (右)	π (左)
状态 1	0.292 806	0.561 901	0.294 316	0.566 653	0.116 791	0.116 180
状态 2	0.190 064	0.042 868	0.190 336	0.041 366	0.321 721	0.323 580
状态 3	0.017 312	0.116 605	0.015 503	0.115 992	0.070 848	0.069 683
状态 4	0.160 129	0.041 269	0.160 040	0.039 749	0.074 966	0.073 851
状态 5	0.185 957	0.146 865	0.186 179	0.146 616	0.213 686	0.214 243
状态 6	0.158 732	0.095 492	0.158 626	0.094 625	0.206 988	0.207 463

表 2 向右和向左的 HMM 部分状态的观察状态概率对比

模型状态	状态 1(右)	状态 2(右)	状态 1(左)	状态 2(左)
观察态 1	1.000	1.000	0.001	0.001
观察态 2	0.001	0.001	0.001	0.001
观察态 3	0.001	0.001	0.001	0.001
观察态 4	0.001	0.001	0.001	0.001
观察态 5	0.001	0.001	1.000	1.000
观察态 6	0.001	0.001	0.001	0.001
观察态 7	0.001	0.001	0.001	0.001
观察态 8	0.001	0.001	0.001	0.001

5.5 实时 HMM 序列的采集并进行识别的实验

HMM 识别手势实时运动轨迹成功时弹出一个确认的对话框,然后开始重新采集为下一次识别做准备。而 HMM 轨迹序列的识别成功与否,要取决于 HMM 序列采集的好坏,在实验过程中发现对左右识别的成功次数比上下识别的成功次

5.3 基于改进的粒子直方图的跟踪实验

如图 10 结合运动检测和人为标定的 HSV 粒子直方图,摄像头能够进行很好的实时跟踪。

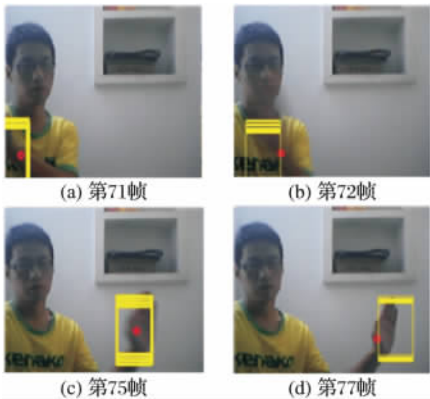


图 10 粒子直方图的手势跟踪

5.4 HMM 训练实验

本文选取的 HMM 状态数为 6,观察状态数为 8,序列长度 T 为 10,训练 HMM 时采用实时时间作为训练 HMM 的随机数种子,且四个 HMM 参数 π A B 都用同一个随机种子产生,因而 HMM 在训练上具有同等条件,由于模型参数的数据众多,这里只列出向左和向右序列的部分模型参数如表 1 和表 2。

从表 1 中可看出对于同一随机种子 2 个模型的状态矩阵 A 部分概率以及初始状态 π 基本一样;而表 2 中,在向右的混淆矩阵中状态 1 和状态 2 在观察状态 1 的概率都为 1.0,在向左的混淆矩阵中状态 1 和状态 2 在观察状态 5 的概率也都为 1.0,而它们的其他观察状态的概率都很小,这说明对向右和向左方向的该状态训练达到了要求,而其他状态以及向上和向下的观察序列训练的 HMM 也都达到了要求,这里没有列出。

数要多,识别错误次数居中,未识别与识别错误相当,如表 3 所示。

表 3 实验数据统计

运动方向	识别成功次数	识别错误次数	未识别次数
向上	5	5	10
向下	7	4	9
向左	11	6	3
向右	14	4	2

程序界面如图 11,从表 3 可以看出,运动轨迹对左右运动的识别率较高,对于实际情况而言,在水平运动方向上没有类肤色的干扰,而垂直运动方向上运动有手臂肤色的干扰,尽管可以穿长袖衣服,但是手部肤色垂直长度要比水平长度要长,且图像尺寸为 320×240 ,因而垂直运动的方向上采集序列的长度要比水平方向运动上的采集序列的长度要短,因而发生未识别的概率增大。



图 11 主程序界面

6 结语

本文针对计算机视觉 结合运动检测、跟踪与识别等算法设计了一个简易实时人机交互系统。在运动检测过程中可以快速准确地检测运动手势区域,说明背景差分的运动检测与直方图的运动检测相结合的方法满足实时性的要求;在跟踪手势过程中,也能及时准确地跟踪手势区域,说明基于直方图的粒子跟踪能满足实时性与运动的不确定性;在识别过程中,能基本完成上下左右轨迹的识别,说明 HMM 轨迹识别能满足识别的有效性。

参考文献:

- [1] 吴晓阳. 基于 OpenCV 的运动目标检测与跟踪[D]. 杭州: 浙江大学, 2008.
- [2] CHO J U, JIN S H, PHAM X D. Object tracking circuit using particle filter with multiple features [C]// SICE-ICASE: International Joint Conference. Las Vegas: IEEE, 2006: 1431 - 1436.
- [3] 孔晓明, 陈一民, 陈养彬, 等. 基于视觉的动态手势识别[J]. 计算机工程与设计, 2005, 26(11): 2934 - 2936.
- [4] 张青苗, 李鹏, 赵勇, 等. 一种基于 OpenCV 实现的运动目标检测算法的研究[J]. 舰船电子工程, 2010, 30(7): 59 - 61.
- [5] QIN WEN, PENG QICONG. An improved particle filter algorithm based on neural network for visual tracking[C]// International Conference on Communications, Circuits and Systems. Las Vegas: IEEE, 2007: 765 - 768.
- [6] 龚翔. 基于粒子滤波的视觉跟踪算法研究[D]. 南京: 南京理工大学, 2009.
- [7] YU JINXIA, LIU WENJING, YANG Y. Improved particle filter algorithms based on partial systematic resampling[C]// IEEE International Conference on Intelligent Computing and Intelligent Systems. Las Vegas: IEEE, 2010: 483 - 487.
- [8] FAZLI S, POUR H M, BOUZARI H. Particle filter based object tracking with sift and color feature[C]// Second International Conference on Machine Vision. Las Vegas: IEEE, 2009: 89 - 93.
- [9] 侯一民, 郭雷, 伦向敏, 等. 运动背景下基于粒子滤波的目标跟踪[J]. 计算机工程与应用, 2007, 43(8): 62 - 64.
- [10] 姚剑敏. 粒子滤波跟踪方法的研究[D]. 长春: 中国科学院长春光学精密机械与物理研究所, 2004.
- [11] YUAN LICH. An improved HMM speech recognition model[C]// International Conference on Audio. Las Vegas: IEEE, 2008: 1311 - 1315.
- [12] TAN WENJUN, WU CHENG DONG, ZHAO SHUYING. Recognition of dynamic hand gesture based on SCHMM model[C]. // Chinese Control and Decision Conference. Log Vegas: IEEE, 2009: 2430 - 2434.

(上接第 127 页)

3 结语

本文针对在当前的视频监控系统中,单一的摄像机难以对一些视场较大的监控区域如小区、仓库、变电站等进行完整覆盖并取得较高分辨率监控图像导致监控效果不佳等问题,提出了一种基于 SIFT 特征匹配算法的异常目标检测方法,即检测出镜头图像与全景图像的匹配特征点,利用得到的转换矩阵将当前图像转换到全景图像,最终通过差分对比得出异常目标的方法。实验结果表明,即使当前图像相比全景图像存在着尺度、亮度、旋转等差异,本方法仍可以在镜头图像里找出与全景图像存在着差异的入侵目标。

参考文献:

- [1] 苗立刚. 视频监控中的图像拼接与合成算法研究[J]. 仪器仪表学报, 2009, 30(4): 857 - 861.
- [2] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International of Computer Vision, 2004, 60(2): 91 - 110.
- [3] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615 - 1630.
- [4] LINDBERG T. Scale-space theory: a basic tool for analyzing structures at different scales [J]. Journal of Applied Statistics, 1994, 21(2): 224 - 270.
- [5] 曹楠, 王萍. 基于 SIFT 特征匹配的图像无缝拼接算法[J]. 计算机与应用化学, 2011, 28(2): 242 - 244.
- [6] 李金, 梁洪. 融合全局特性的 SIFT 特征在图像检索中的应用[J]. 黑龙江大学学报: 理工 C 版, 2010, 1(1): 92 - 95.
- [7] 严磊, 汪增福. 基于多局部特征匹配的全自动图像拼接[J]. 计算机应用与软件, 2010, 27(10): 5 - 7.
- [8] BEIS J S, LOWE D G. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces[C]// Proceedings of the IEEE 1997 Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 1997: 1000 - 1006.
- [9] 王宇, 王涌天, 刘越. 基于 SIFT 和小波变换的图像拼接算法[J]. 北京理工大学学报, 2009, 29(5): 423 - 226.
- [10] 范旭丰. 室内环境下基于单目视觉的目标跟踪方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2008.
- [11] SHUM H Y, SZELISKI R. Construction of panoramic image mosaics with global and local alignment[J]. International Journal of Computer Vision, 2000, 36(2): 101 - 130.
- [12] ZOGLAMI I, FAUGERAS O, DERICHE R. Using geometric corners to build a 2D mosaic from a set of images[C]// Proceedings of the International Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 1997: 167 - 169.