

基于 Camshift 反馈码本模型的手抓取物跟踪方法

赵翠莲 王 洪 范志坚 郭 晶

(上海大学机电工程与自动化学院上海市智能制造及机器人重点实验室 上海 200072)

摘 要 针对 Camshift 算法在复杂背景下无法自动跟踪目标的问题,提出一种基于 Camshift 反馈码本模型的运动目标检测和跟踪算法。该算法首先利用码本模型检测前景目标,然后采用 Camshift 在颜色概率分布图中跟踪前景区域中的目标,通过窗口尺寸比较和直方图相关性判断来解决自动跟踪,通过窗口位置预测和尺寸扩大来改进下一帧 Camshift 算法的输入搜索窗口,同时并集操作多个目标处理后的矩形窗口,并将其反馈为下一帧码本模型的图像检测区域。最后将该算法应用于手和目标物的抓取状态判断上,具体过程是在静态背景下利用两个摄像头采集到的图像进行手和目标物的检测和跟踪,通过矩形相交性判断抓取次数,以验证跟踪算法的有效性。实验结果表明,通过信息反馈减小了目标检测和跟踪的搜索区域,提高了算法的实时性,在单摄像头下可提高处理帧频 130%。

关键词 Camshift, 码本模型, 跟踪, 抓取物

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.12.064

Tracking Method of Hand Grasping Objects Based on Feedback from Camshift to Codebook Model

ZHAO Cui-lian WANG Hong FAN Zhi-jian GUO Jing

(Shanghai Key Laboratory of Intelligent Manufacturing and Robotics, School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China)

Abstract In order to solve problem that Camshift algorithm can't automatically track targets in complex background, an algorithm to detect and track moving objects based on feedback from Camshift to codebook model was proposed. At first, the algorithm detects foreground objects by codebook model and tracks objects in the foreground area by Camshift using color probability distribution. Automatic tracking is achieved by comparison of window sizes and judgment of histogram correlation, and an input search window of the next frame is improved by location prediction and size expansion of window. At the same time, a union of multi-processed rectangular windows is as a feedback to the next frame of the image detection region for codebook model. Finally, the algorithm is applied for determining a grasping status between a hand and objects. The specific process is using the captured images by two cameras to detect and track hand and objects in a static background and then counting the number of grasp by rectangle intersection to verify the tracking algorithm. Since the information feedback reduces the search region of detection and tracking, this algorithm gets a better real-time performance, and some experimental results show that the frame processing rate can increase 130% in a single camera.

Keywords Camshift, Codebook model, Tracking, Grasping objects

在计算机视觉研究领域,运动目标检测和跟踪技术已成为一个重要的研究方向。目前在静态背景下精确跟踪如人手这样的非刚性物体依然相当困难,一方面因为人手形状、肤色深浅各异,人手和手指可以移动很快,在几帧之间变化很大^[1],另一方面人手是与自然与社会进行交互的重要器官,其运动过程往往触及生活中的用品和场景。然而由于人手运动过程蕴含着人的病理和生理等特征信息,如利用手与物进行抓握动作时的视觉信息分析帕金森患者的病理并评估其康复情况^[2],因此近年来,人手和抓取物的多目标检测和跟踪技术在智能感知与人机交互^[3]、虚拟现实^[4]、临床医学和康复工程^[5]等领域正受到广泛的关注。

Camshift(Continuously Adaptive mean-shift)算法是在 Mean-shift 算法上的改进并克服了 Mean-shift 算法的固定窗

口的缺点^[6]。Camshift 算法利用选定矩形窗口内目标的颜色直方图计算每帧图像的概率分布,并根据上一帧的结果自适应调整搜索窗口的位置和大小,迭代计算到密度最大的区域,得到当前帧中目标的矩形窗口。CamShift 作为一种非常高效的算法被广泛用于跟踪各种形状特征的对象。

但是跟踪的高效性需要有好的目标检测方法。倘若由于较差的目标检测使其初始跟踪窗口含有部分背景,将会导致跟踪数据的不收敛和丢失。本质上说,CamShift 技术形成的跟踪窗口是跟踪概率分布的峰值,因此在复杂场景下提高 Camshift 算法的稳定性已成为近年来研究的热点。

所提方法主要是将 Camshift 算法与其他一些跟踪算法进行组合以改善其性能。文献[7]提出一种将双向差分相乘与基于仿射运动模型的匹配差分相结合的运动目标检测方法

到稿日期:2014-12-14 返修日期:2015-04-20

赵翠莲(1963—),女,博士,教授,主要研究方向为图像检测、生机电一体化技术;王 洪(1989—),男,硕士生,主要研究方向为图像跟踪;范志坚(1982—),男,硕士,工程师,主要研究方向为生机电工程;郭 晶(1989—),男,硕士生,主要研究方向为信息融合与软硬件集成。

法,文献[8]在颜色概率分布图计算中通过在帧间差分法加入目标运动信息,自动初始化跟踪并排除与目标相似背景颜色的干扰。许多学者也考虑到预测模型可以改进 Camshift 算法的收敛速度,文献[9,10]结合 Camshift 和 Kalman 滤波解决当遮挡发生时 Camshift 算法目标跟踪丢失的问题。

码本模型(codebook)是 Kim 等人提出的一种运动目标检测方法^[11,12]。为了有效地确定跟踪目标,减小背景或其他相似颜色物体的干扰,本文提出采用码本模型进行目标检测,在传统 Camshift 算法的基础上,通过比较窗口尺寸,计算直方图相关性,对下一帧目标搜索窗口位置进行运动估计,改进 Camshift 算法,并利用上述改进 Camshift 算法的输出作为前景码本模型的反馈输入,以减少计算量,提高目标跟踪算法的鲁棒性和准确性。最后将该跟踪算法应用在静态场景下手和目标物的抓取次数判定上,利用两个摄像头采集到的图像进行手和目标物的检测和跟踪,通过矩形相交性判断抓取次数,以验证跟踪算法的有效性。

1 Camshift 反馈码本模型的算法

1.1 算法的描述

设 $F = \{f_1, \dots, f_n\}$ 是一个实时视频序列,在满足实时性要求的情况下,在运动目标的检测和跟踪中,一般先检测后跟踪,检测的结果作为跟踪算法的输入。但是由于目标的运动位置存在连续性,而处理算法是在每一帧频中不断地检测和跟踪循环的,为了提高跟踪算法的实时性,减少处理的信息量,本文提出利用跟踪的结果作为下一帧检测的对象,即将改进 Camshift 算法的输出作为前景码本模型的反馈输入的跟踪算法。其基本流程如图 1 所示。

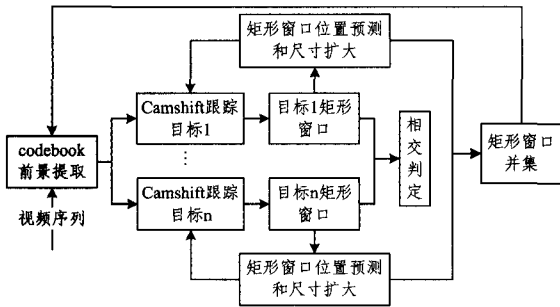


图 1 反馈算法执行流程

采用并重复如下步骤进行跟踪计算,直到程序退出。

- 1) 读入一帧视频图像,在矩形窗口区域上利用码本模型提取前景运动目标,排除背景颜色的干扰;
- 2) 利用 Camshift 算法跟踪前景区域中的目标,得到目标的矩形窗口;
- 3) 矩形窗口的位置利用运动估计预测和尺寸扩大,作为下一帧目标跟踪的搜索窗口;
- 4) 把所有目标位置预测和尺寸扩大之后矩形窗口的并集反馈为下一帧码本模型前景提取的图像处理区域。

每次在对运动目标进行跟踪之前,需先预处理跟踪目标。首先从文件中加载代表跟踪目标的图像和矩形数据,初始化 Camshift 算法。

1.2 codebook 前景目标提取

codebook 采用一种聚类技术,从背景图像序列中建立一个背景模型,它为每个像素建立一个码本(codebook, CB),每个码本由一个或多个码字组成(codeword, CW),并根据亮

度界限范围(box)采样每个像素,聚合成一组码字。

假设 $X = \{x_1, x_2, \dots, x_N\}$ 是一个像素的训练视频序列,其中 $x_t (t=1, \dots, N)$ 是 YUV 向量,其中“Y”表示明亮度(Luminance 或 Luma),也就是灰阶值;而“U”和“V”表示色度(Chrominance 或 Chroma)。并设 $C = \{c_1, c_2, \dots, c_L\}$ 表示组成该像素码本的 L 个码字,每个码字 $c_i (i=1, \dots, L)$ 定义为三元组结构。

$$b_i^{\min} = (Y_i, U_i, V_i), b_i^{\max} = (Y_i, U_i, V_i)$$

$$l_i^{\min} = (Y_i, U_i, V_i), l_i^{\max} = (Y_i, U_i, V_i)$$

$$u_i = \langle f_i, s_i \rangle$$

式中, b_i^{\min}, b_i^{\max} 是 YUV 颜色空间中亮度界限范围 box 的分界线; l_i^{\min}, l_i^{\max} 是学习的界限,触发产生一个新的码字的阈值范围, f 表示码字出现后最后一次匹配时间,在此设为帧序数; s 表示在训练更新过程中码字最后一次更新到当前时间的最大间隔。

在构建码本的训练期间,将 t 时刻的像素点 x_t 的采样值与当前的码本做比较,如果有码字 $c_m (m$ 是某个码字的索引)与其匹配,则以码字 c_m 作为该采样点的编码近似值,并更新其属性,以下是码本提取的算法步骤。

1. $L \leftarrow 0, C \leftarrow \emptyset$
2. for $t = 1$ to N do
 - (1) $x_t = (Y, U, V), l_t = x_t - cb, h_t = x_t + cb$
 - (2) 在集合 $C = \{c_i | 1 \leq i \leq L\}$ 中寻找与 x_t 匹配的码字 c_m , 条件是

$$l_m^{\min} \leq x_t \leq l_m^{\max}$$
 - (3) 如果 $C = \emptyset$ 或者没有找到匹配,则 $L \leftarrow L + 1$ 并创建一个新的码字 $c_L: (b_L^{\min}, b_L^{\max}) \leftarrow (x_t, x_t), (l_L^{\min}, l_L^{\max}) \leftarrow (l_t, h_t), u_L \leftarrow \langle t, 0 \rangle$
 - (4) 否则更新匹配的码字 c_m 结构,设置如 $(b_m^{\min}, b_m^{\max}) \leftarrow (\min(b_m^{\min}, x_t), \max(b_m^{\max}, x_t)),$ if $l_m^{\min} > l_t$ then $l_m^{\min} \leftarrow l_m^{\min} - 1,$ if $l_m^{\max} < h_t$ then $l_m^{\max} \leftarrow l_m^{\max} + 1, u_m \leftarrow \langle t, 0 \rangle$
 - (5) for $i = 1$ to L do

$$s_i \leftarrow \max(s_i, t - f_i)$$
 end for
3. for $i = 1$ to L do
 - if $s_i \leq T_M$ then 删除 c_i

其中阈值 cb 表示背景波动的范围,阈值 T_M 通常取训练帧数的一半即 $N/2$,表示所有代表背景的码字必须至少在 $N/2$ 帧中出现。训练过程中得到的码本存在冗余,其中可能有一些表示前景运动目标和噪声的码字,引入时间间隔变量 s ,利用步骤 3 即可将这些码字从概率意义上分离出去,从而允许初始训练过程中存在运动目标。

当创建好背景码本模型后,码本模型根据像素采样值与其码本是否匹配来判断前景或背景,对于运动目标检测过程中的新输入像素 $x = (Y, U, V)$ 及其对应的码本 M ,前景提取操作 $BGS(x)$ ^[13] 计算步骤如下。

1. $x = (Y, U, V), l = x - m^{\min}, h = x + m^{\max}$
2. $match \leftarrow 0,$
 - for $i = 1$ to L do
 - if $(b_i^{\min} \leq l) \text{ and } (h \leq b_i^{\max})$ then $match \leftarrow 1$

$$3. BGS(x) = \begin{cases} x \text{ 为前景像素,} & match = 0 \\ x \text{ 为背景像素,} & match = 1 \end{cases}$$

即像素大小在某个码字的分界范围内,判断为背景像素,否则认为是前景像素。阈值 m^{\min}, m^{\max} 是背景像素值的上下界误

差,需针对不同的光照环境调整大小,以保证最好的提取效果。

1.3 传统的 Camshift 算法

Camshift 算法通过目标的颜色直方图把输入图像转化为概率分布图,在转换的概率分布图中计算目标区域的各阶矩,通过对获取的目标矩形窗口的位置和尺寸大小连续迭代计算,并把矩形窗口扩大后作为下一帧的图像处理区域,实现连续跟踪。传统的 Camshift 目标跟踪步骤为:

- 1) 初始化跟踪物体矩形区域;
- 2) 提取每帧图像 HSV (Hue, 色调; Saturation, 饱和度; Value, 亮度) 颜色空间中 H 分量图像, 计算窗口的重心位置 (x_c, y_c) ;
- 3) 将矩形窗口的中心移动到重心, 并更新矩形窗口;
- 4) 返回目标的矩形窗口位置和大小。

如果移动距离大于收敛的最小移动距离或迭代次数小于最大迭代次数, 则重复步骤 3)、步骤 4), 直到收敛为止。

2 改进 Camshift 算法

2.1 结合窗口比较和直方图相关性判断的自动跟踪

直接采用上述传统的 Camshift 算法进行目标跟踪时, 由于目标从摄像头视野中离开或者被完全遮挡, 跟踪目标丢失, 目标矩形窗口尺寸为 0, 使得系统出现崩溃; 或者由于噪声的影响跟踪窗口变成很大的区域, 致使矩形窗口与目标没有联系, 即使当目标再次出现时, 也无法自动恢复跟踪。出现上述原因是由于传统的 Camshift 算法对搜索到的矩形窗口尺寸和目标之间的相关性缺少判断。为此本文提出改进的 Camshift 算法流程如图 2 所示。

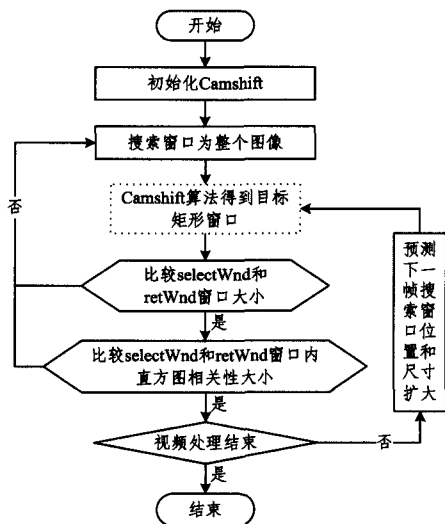


图 2 改进的 Camshift 算法流程

该改进算法在传统 Camshift 算法的基础上, 通过比较目标矩形窗口与初始化矩形窗口的尺寸大小, 排除矩形窗口尺寸为 0 和输入噪声的影响, 并联合直方图相关性判断实现自动检测与跟踪, 防止跟踪目标丢失。

首先, 在对运动目标进行跟踪之前, 从文件中读取目标图像和窗口数据来初始化 Camshift 算法, 计算图像 HSV 颜色空间中 H 分量在矩形内的颜色直方图; 其次, 设置搜索窗口为整个图像尺寸, 使得第一帧时 Camshift 在整个图像区域上搜索目标。

在跟踪过程中, 加入目标跟踪丢失判断, 并自适应调整 Camshift 算法的输入和输出窗口。令 $searchWnd$ 为输入窗口, $retWnd$ 为输出窗口, $selectWnd$ 为文件中读取的窗口数据。比较 $retWnd$ 与 $selectWnd$ 的尺寸, 因为跟踪的窗口大小在帧间变化幅度有限, 在此可设 $\frac{1}{4} \times selectWnd < retWnd < 4 \times selectWnd$ 表示搜索到目标, 否则, 重新设置搜索区域为整个图像尺寸, 进行下一帧搜索。对满足尺寸大小的 $retWnd$ 计算采用式(1)计算直方图相关性, 进一步确认窗口内的目标是否为跟踪目标。

$$d(H_1, H_2) = \sum_I \frac{(H_1(I) - H_2(I))^2}{H_1(I) + H_2(I)} \quad (1)$$

式(1)计算的结果越小, 相似度越大。以初始化时矩形窗口内 H 分量的直方图为准, 计算每帧目标矩形窗口 H 分量的直方图与基准的相关性值 d , 设置阈值 t , 当 $d < t$ 时, 表示矩形窗口内的目标为跟踪目标; 否则, 重新设置搜索区域为整个图像尺寸, 进行下一帧搜索。对于满足条件的目标矩形窗口, 利用运动估计进行位置预测和尺寸扩大, 以作为下一帧 Camshift 算法的输入搜索窗口。

2.2 运动估计

当目标高速运动时, 若 Camshift 算法根据当前图像跟踪结果得到的矩形窗口与上一帧矩形窗口之间的位移过大, 两个窗口之间没有重叠, 会导致跟踪丢失。为了增大矩形窗口与目标联接的相关性, 提高跟踪的稳定性, 在运动目标跟踪过程中, 可以对下一帧目标搜索窗口的位置进行简单的运动估计, 利用得到的预测值对窗口进行位置平移和尺寸扩大。预测目标的运动参数的方法有 Kalman 滤波^[9], 但本文通过实验发现, 当目标在一个方向上高速运动并突然反向时, Kalman 预测的矩形窗口与实际的目标矩形窗口没有重叠, 导致目标跟踪丢失, 而本文提出的线性预测的计算只与前一帧有关, 即使发生如上情况, 预测的矩形窗口也能与目标实际的矩形窗口快速地匹配, 因此, 从跟踪效果上考虑, 本文采用简单的线性预测。

由于时间很短, 假设目标在相邻两帧内可以看成做匀速直线运动。令第 i 帧到第 $i-1$ 帧之间目标在水平方向上的移动距离为 Δx_{i-1} , 垂直方向上移动距离为 Δy_{i-1} , 时间间隔为 t_{i-1} 。同样, t_i 为第 i 帧到第 $i+1$ 帧之间的时间间隔, 预测目标在第 $i+1$ 帧的水平移动距离为 Δx_i^* , 垂直移动距离为 Δy_i^* , 令:

$$\Delta x_i^* = \frac{\Delta x_{i-1}}{t_{i-1}} t_i; \Delta y_i^* = \frac{\Delta y_{i-1}}{t_{i-1}} t_i \quad (2)$$

然后将代表目标的矩形窗口位置按 Δx_i^* 和 Δy_i^* 进行平移, 由于目标的随机性运动, 特别是当多个目标运动轨迹交叉发生遮挡时, 即使采用如上方法对矩形窗口进行位置预测, 两帧之间的矩形窗口仍有可能不相交, 进而导致跟踪丢失, 因此, 为了提高算法的稳健性, 需要根据位移大小将预测矩形窗口尺寸扩大, 作为下一帧图像的搜索区域。

$$\begin{cases} searchWnd.x = \min(retWnd.x, retWnd.x + \Delta x_i^*) \\ searchWnd.y = \min(retWnd.y, retWnd.y + \Delta y_i^*) \\ searchWnd.width = retWnd.width + 2|\Delta x_i^*| \\ searchWnd.height = retWnd.height + 2|\Delta y_i^*| \end{cases} \quad (3)$$

式(3)表示扩大的矩形窗口为跟踪结果窗口和预测搜索窗口的并集, 因此是增加 2 倍的关系, 但是没有考虑到目标在运动

过程中变形而导致矩形窗口大小发生变化的情况,当目标运动缓慢时,搜索的窗口会逐步退化,把原是目标的像素区域剔除,因此,需要加入一定的扩大区域。

$$\begin{cases} searchWnd.x=searchWnd.x-region \\ searchWnd.y=searchWnd.y-region \\ searchWnd.width=searchWnd.width+2*region \\ searchWnd.height=searchWnd.height+2*region \end{cases} \quad (4)$$

式中, $region$ 为扩大尺寸的阈值,取值大小与运动目标的速度成正比,通常,对于高速运动的目标,取值较大;反之,取值较小。然而,该值的大小还影响 Camshift 算法输入的搜索窗口,值越大搜索的区域就越大,计算量越多,但是获得的跟踪结果更加准确;反之,计算量越小,但是目标跟踪丢失的概率增大。综上所述,需要在算法速度和稳定性上作出一个折中,可以通过实验验证取值大小,以获得更好的跟踪稳定性。

通过以上对算法的改进,不管目标在摄像头视野内完全遮挡后重新出现,还是重新进入到视野内,算法都能自动检测并跟踪目标。

3 基于双目的抓取实验与结果分析

3.1 双目实验平台的搭建

临床科学中的上肢动作研究测试理论 (Action Research Arm Test, ARAT)^[14] 是针对手部运动功能的测试评估方法之一,针对手部功能障碍患者在康复训练过程中手抓取目标物数字化的问题,搭建的实验平台如图 3 所示。

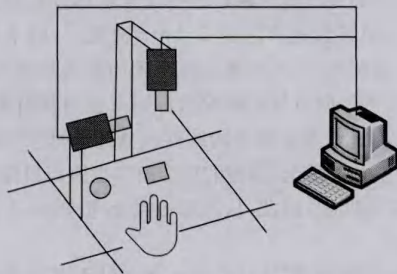


图3 双目实验平台示意图

两摄像头成 90° 摆放,这样可以获取两个正交平面内的图像数据。两个摄像头的图像通过 USB 传输到 PC 上,并利用 DirectShow 实现同步采集。

利用文中算法计算手和目标物的矩形跟踪窗口,通过矩形相交性算法(见图 1)判断手和目标物的抓取次数,即为了提高判断的准确性,对两个垂直摄像头同步采集到的图像相交性结果作与运算,只有在两张图像内的手和目标物矩形都相交的情况下,才判定手抓取到目标物。

3.2 码本模型前景目标提取实验

利用背景图像序列建立好码本模型后,对输入的每帧图像进行 $BGS(x)$ 操作来提取前景目标。在静态背景下,放入场景中的跟踪目标在光照情况下会产生阴影,从而降低了前景目标提取的准确度。在实验时手动调节 m^{min} 、 m^{max} 的值,剔除阴影对前景目标提取的影响。

设置背景波动的范围 $cb=(10,10,10)$ 并建立码本模型,如图 4 所示为原始画面在不同 m^{min} 、 m^{max} 值下的提取效果。在图 4(b)中可以看到光照产生的阴影对提取效果有很大的影响;图 4(c)中增大了阈值最小值 m^{min} ,剔除了阴影的影响;

图 4(d)中调整阈值最大值 m^{max} ,把原是方块的像素也剔除了。在相同环境下,只需要训练一次,目测效果对目标物取图 4(c)的值。

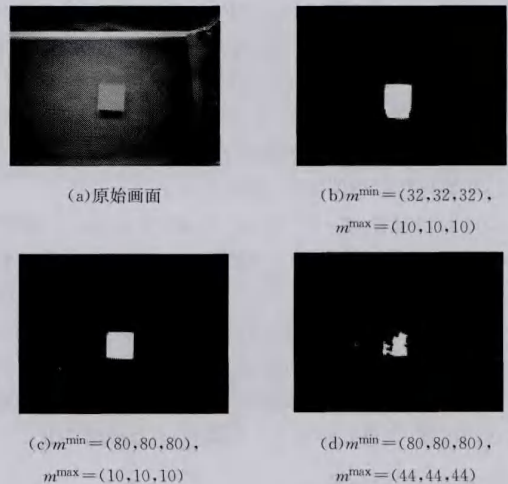


图4 前景目标提取

3.3 改进 Camshift 算法的实验

把手抓取方块两次的连接视频图像作为实验对象,相关性系数 d 设置为 20000,扩大尺寸的阈值 $region$ 为 20 pixels。

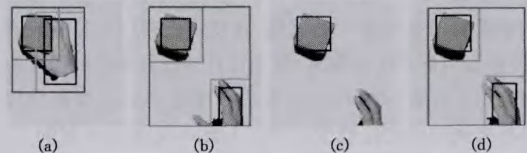


图5 自动检测与跟踪

实验结果如图 5 所示,图中黑色矩形框为跟踪结果,比黑色矩形框稍大的是位置预测和尺寸扩大的矩形框,作为下一帧 Camshift 算法的输入窗口。两个矩形框的并集(最外层的矩形框)作为下一帧码本模型的感兴趣区域。其中图 5(a)为手和方块的矩形框相交,判定手抓取到方块,图 5(b)为手慢慢地离开摄像头视野,图 5(c)为手矩形框尺寸小于阈值时,自动判断跟踪丢失,图 5(d)为当目标再次出现在摄像头视野中时,自动检测到手并继续跟踪。

3.4 Camshift 算法反馈和不反馈码本模型性能对比实验

为了对比 Camshift 反馈和不反馈码本模型情况下系统运行的实时性,分别以单摄像头和双摄像头进行实验,当检测和跟踪运行稳定时,处理帧频如表 1(单位为 fps,即每秒处理的帧数)所列。

表1 反馈和不反馈条件下的帧率(fps)

摄像头个数	无反馈矩形框	反馈矩形框
单摄像头	28	32
双摄像头	16	21

从表 1 记录的处理帧频可见,基于反馈的算法由于只在感兴趣区域上处理图像,减少了计算量,不论在单摄像头还是双摄像头的情况下,都比没有输入反馈的条件下实时性好。在单摄像头下可提高处理帧频 130%。

3.5 抓取次数判定流程

实验中,训练者在摄像头视野内抓取方块,抓取次数判断流程如图 6 所示。

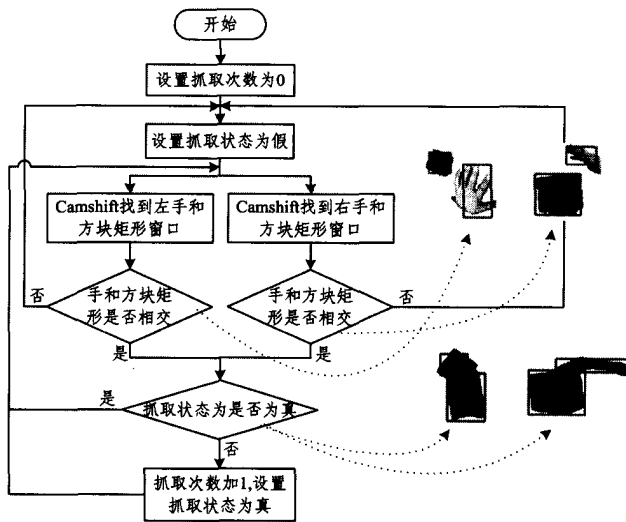


图6 手和抓取物相交判定流程

开始时,设置抓取次数为0,为了保证对抓取不重复计数,引入抓取状态变量,利用 Camshift 算法分别对两个摄像头得到的图像进行跟踪,找到手和抓取物的矩形窗口,通过矩形框相交性判断手是否抓取到方块,计算公式如式(5)所示:

$$\begin{cases} |X_{b2} + X_{b1} - X_{a2} - X_{a1}| \leq X_{a2} - X_{a1} + X_{b2} - X_{b1} \\ |Y_{b2} + Y_{b1} - Y_{a2} - Y_{a1}| \leq Y_{a2} - Y_{a1} + Y_{b2} - Y_{b1} \end{cases} \quad (5)$$

式中, (X_{a1}, Y_{a1}) 表示手矩形窗口的左上角坐标, (X_{a2}, Y_{a2}) 表示手矩形窗口的右下角坐标, 矩形 B 的左上角坐标为 (X_{b1}, Y_{b1}) , 表示方块矩形窗口左上角坐标, (X_{b2}, Y_{b2}) 表示右下角坐标。只有满足两个摄像头图像平面内的手和方块矩形窗口都相交且抓取状态为假时,设置抓取次数加1,此时,设置抓取状态为真,防止重复计数,在其他任何状态下,如果 Camshift 没有跟踪到目标或只有一个摄像头采集的图像中手和方块矩形窗口相交,都设置抓取状态为假,进行下一帧判断。

手和方块在跟踪过程中会发生形变、缩放、旋转等,如果直接采用本文算法计算得到的矩形进行相交性判断,会不精确,可以对矩形框缩小一定区域后再进行相交性判断,提高判断的准确性。对于抓取次数为10的同一视频图像序列,矩形区域的缩小与相应判断的次数关系如图7所示。

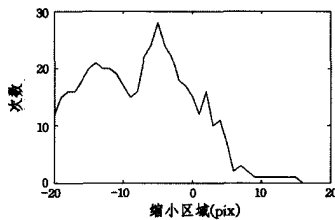


图7 判断次数与区域缩小的关系

从图中可以看出,当缩小区域取3个像素时,抓取次数判断为实际抓取次数10次,在取其他值时,误差比较大。因此,每次训练时需要通过实验预设置矩形区域缩小值,以得到抓取次数更精确的判断值。

结束语 本文提出了一种基于 Camshift 反馈码本模型的检测和跟踪算法。该算法在前景区域中对单个目标利用 Camshift 跟踪,通过窗口尺寸比较、直方图相关性计算、运动估计把 Camshift 跟踪得到的窗口矩形位置进行预测和尺寸扩大,以作为下一帧跟踪的输入搜索窗口,把多个目标矩形窗口处理之后的并集作为下一帧码本模型的图像检测区域,减少了计算量,提高了算法效率。通过实验证明了该算法的稳

定性。实验结果表明,通过反馈限定目标检测和跟踪的区域可以提高算法的实时性,因为该反馈算法在单摄像头下可提高处理帧频130%。本文对两个垂直摄像头拍摄到的图像平面内手和目标物矩形窗口的相交性进行判断是否抓取到,但对于手包围目标物但未接触到目标物的情况无法处理,在未来工作中将研究通过立体视觉获得手和目标物的空间位置和旋转信息,以及利用多传感信息进行抓握交互的理论研究。

参考文献

- [1] Rautaray S S, Agrawal A. Vision based hand gesture recognition for human computer interaction; a survey[J]. Artificial Intelligence Review, 2012; 1-54
- [2] Arias P, Robles-García V, Sanmartín G, et al. Virtual reality as a tool for evaluation of repetitive rhythmic movements in the elderly and Parkinson's disease patients[J]. EE Times Asia, PLoS One, 2012, 7(1); e30021
- [3] Bradski G R. Computer vision face tracking for use in a perceptual user interface[J]. EE Time Asia, 1998, 2(2); 1-15
- [4] Varona J, Buades J M, Perales F J. Hands and face tracking for VR applications[J]. Computers & Graphics, 2005, 29(2); 179-187
- [5] Sanmartín G, Flores J, Arias P, et al. Motion capture for clinical purposes, an approach using primesense sensors[M]// Articulated Motion and Deformable Objects. Springer Berlin Heidelberg, 2012; 273-281
- [6] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5); 603-619
- [7] 李劲菊, 朱青, 王耀南. 一种复杂背景下运动目标检测与跟踪方法[J]. 仪器仪表学报, 2010, 31(10); 2242-2247
Li J J, Zhu Q, Wang Y N. Detecting and tracking method of moving target in complex environment [J]. Chinese Journal of Scientific Instrument, 2010, 31(10); 2242-2247
- [8] 邬大鹏, 程卫平, 于盛林. 基于帧间差分法和运动估计的 Camshift 目标跟踪算法[J]. 光电工程, 2010, 37(1); 55-60
Wu D P, Cheng W P, Yu S L. Camshift Object Tracking Algorithm Based on Inter-frame Difference and Motion Prediction [J]. Opto-Electronic Engineering, 2010, 37(1); 55-60
- [9] Huang S, Hong J. Moving object tracking system based on camshift and Kalman filter[C]// 2011 International Conference on Consumer Electronics, Communications and Networks (CEC-Net). IEEE, 2011; 1423-1426
- [10] Weia Z, Zhangb L, Kima H, et al. An Improved Object Tracking Algorithm Based on Camshift Combined with Active Contour and Kalman Filter[J]. Journal of Information & Computational Science, 2014, 11(6); 1753-1764
- [11] Kim K, Chalidabhongse T H, Harwood D, et al. Background modeling and subtraction by codebook construction[C] // 2004 International Conference on Image Processing, 2004 (ICIP'04). IEEE, 2004; 3061-3064
- [12] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time foreground-background segmentation using codebook model[J]. Real-time Imaging, 2005, 11(3); 172-185
- [13] 张军, 代科学, 李国辉. 基于 HSV 颜色空间和码本模型的运动目标检测[J]. 系统工程与电子技术, 2008, 30(3); 423-427
Zhang J, Dai K X, Li G H. HSV color-space and codebook model based moving objects detection [J]. Systems Engineering and Electronics, 2008, 30(3); 423-427
- [14] Yozbatiran N, Der-Yeghian L, Cramer S C. A standardized approach to performing the action research arm test[J]. Neurorehabilitation and Neural Repair, 2008, 22(1); 78-90