

基于 SIFT 和感知哈希改进的 CamShift 跟踪算法^{*}

李文举 王子杰 崔柳

(上海应用技术大学计算机科学与信息工程学院 上海 201418)

摘要: 传统的 CamShift 仅使用目标的颜色直方图作为特征,因此在相似背景干扰、遮挡、高速运动等情况下容易出现跟踪不准确或丢失跟踪目标的现象。针对上述不足,提出了基于 SIFT 和感知哈希改进的 CamShift 跟踪算法。首先,将图像从 RGB 颜色空间转为 HSV 颜色空间,分别出色调和饱和度直方图,并提取图像的边缘梯度直方图进行融合获得目标的融合直方图。其次在 CamShift 算法框架下得到最优候选目标,若候选目标与目标模板的 Bhattacharyya 距离大于阈值时,则使用改进的感知哈希算法进行最优候选目标的搜索。然后在下一帧搜索时,在目标和视频序列的高信息熵部分使用 SIFT 算法进行特征点的提取并匹配从而获取初始搜索窗口,若 SIFT 算法匹配失败,则使用卡尔曼滤波预测的搜索框作为初始搜索窗口。将该算法首先在 OTB-100 数据集上和其他常用的跟踪算法进行对比实验,实验结果表明算法能够准确地跟踪目标,跟踪成功率达到了 90.1%。将该算法应用于人脸跟踪任务中,并与其他的人脸跟踪算法进行对比实验,实验结果表明该算法具有更好的性能和准确性,跟踪成功率达到了 93.5%。

关键词: 目标跟踪;感知哈希;SIFT;CamShift;卡尔曼滤波;直方图;人脸跟踪

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Improved CamShift tracking algorithm based on SIFT and perceptual hash

Li Wenju Wang Zijie Cui Liu

(School of Computer Science and Information Engineering, Shanghai Institute of Technology, Shanghai 201418, China)

Abstract: The traditional CamShift only uses the color histogram of the target as the feature, so it may lead to inaccurate tracking or losing the target in the case of similar background, occlusion, high-speed motion and so on. In view of the above shortcomings, an improved CamShift tracking algorithm based on SIFT and perceptual hash is proposed. Firstly, transforming the image from RGB color space to HSV color space, then extracting the hue and saturation histograms and the edge gradient histogram of the image and combine the histograms to obtain the fusion histogram of the target. Secondly, using the fusion histogram of the target to obtain the optimal candidate target under the framework of CamShift algorithm. If the Bhattacharyya distance between the candidate target and the target template is larger than the threshold, using the improved perceptual hash algorithm to search the optimal candidate target. Then in the next frame search, using the SIFT algorithm to extract the feature points of the high information entropy part of both the target and video sequence, then matching the feature points to obtain the initial search window. If the SIFT algorithm fails to match, using the search box which is predicted by the Kalman filter as the initial search window to search the target. The algorithm is compared with other common tracking algorithms on OTB-100 dataset. The experimental results show that the algorithm can track the target accurately and the success rate reaches 90.1%. Then applying the algorithm to the task of face tracking and compared with other face tracking algorithms. The experimental results show that the algorithm has good performance and high accuracy, and the tracking success rate reaches 93.5%.

Keywords: target tracking; perceptual hashing; SIFT; CamShift; Kalman filtering; histogram; face tracking

0 引言

目标跟踪是计算机视觉中重要的研究方向之一,目标

跟踪在实际生活中有着广泛的应用,例如智能小车导航^[1]、水下目标跟踪^[2]和车辆跟踪^[3]等。目标跟踪的主要任务是在连续的动态视频帧中准确找到目标的位置,然而要达到

收稿日期:2022-06-30

^{*} 基金项目:国家自然科学基金(61973307,61903256)项目资助

这一目标是具有一定的困难的,因为在实际的视频序列中存在着很多不确定的干扰因素,其中一些干扰因素是由跟踪目标本身引起的,如形变、旋转、位置变化或运动速度突变等;其它一些干扰因素是由外部环境变化所造成的,如光照变化和背景杂波等。所有上述的干扰因素都会使得跟踪目标的视觉特性发生改变,从而导致跟踪精度降低。

传统的 CamShift 跟踪算法以跟踪目标的颜色作为特征建立颜色直方图,通过搜索和目标颜色特征最佳匹配区域来实现对目标的跟踪, CamShift 跟踪算法能够解决遮挡、形变等问题,并且因为其计算量小的优点,在目标跟踪领域有着广泛应用,并在人脸跟踪^[4-7]和无人机目标跟踪^[8-9]等实际场景中得到了应用。然而因为 CamShift 跟踪算法仅利用跟踪目标的颜色作为特征,因此在相似背景或相似物遮挡的情况下很容易受到影响,而发生跟踪不准确或丢失目标的现象。鉴于上述问题,国内外很多学者对 CamShift 跟踪算法进行了改进。文献[10]利用 HSV 空间中 H 和 S 分量的组合建立二维颜色特征直方图并结合图像的 LBP 特征直方图,提高了跟踪精度。文献[11-12]使用色度和边缘加权联合直方图表征目标,有效解决了相似色干扰、遮挡等问题。文献[10-12]中的算法均对跟踪目标的直方图进行了改进从而获取了更丰富的目标信息,但并未考虑当目标的外观模型发生变化,从而使得直方图信息不准确导致跟踪精度降低的问题。文献[13]利用 SURF 算法对目标模板和视频序列进行特征匹配,将匹配成功的运动目标的跟踪窗口作为 CamShift 算法的初始搜索窗口,提高了跟踪的稳定性。但该算法并未考虑到当匹配的特征点数量过少时从而导致 SURF 算法匹配失败的情况。文献[14]和[15]利用目标的 LBP 特征,并将其与 CamShift 算法相结合,提升了目标在相似背景干扰下的跟踪准确率。但是文献[14]与[15]中的方法同样未考虑到当目标的外观模型变化过大时, LBP 特征可能不准确从而导致跟踪精度下降的问题。

基于上述文献的研究成果与不足之处,本文提出了基于 SIFT 和感知哈希改进的 CamShift 跟踪算法。首先,将目标从 RGB 颜色空间转为 HSV 颜色空间,分别得出色调和饱和度直方图,并提取图像的边缘梯度直方图进行融合获得目标的融合直方图。然后在 CamShift 算法框架下得到最优候选目标,若候选目标与目标模板的 Bhattacharyya 距离大于阈值时,则使用改进的感知哈希算法进行最优候选目标的搜索。感知哈希算法能够很好地提取跟踪目标的主体特征,本文在感知哈希的基础之上进行改进从而提取更丰富的目标特征。在下一帧搜索时,对图像的高信息熵部分使用 SIFT 算法进行特征点的提取从而达到加速特征点提取的效果,并对特征点进行匹配从而获取初始搜索窗口。当目标的外观模型变化过大时可能会导致 SIFT 算法匹配的特征点数量过少,从而导致匹配失败。因此本文使用卡尔曼滤波对目标的运动进行预测,当 SIFT 算法匹配

失败时,则使用卡尔曼滤波预测的搜索框作为初始搜索窗口。将本文算法和其他常用的目标跟踪算法首先在 OTB-100 数据集上进行对比实验,实验结果证明,本文算法在目标受到相似背景干扰、遮挡和光照变化时,相比其他算法有着较高的跟踪准确率。然后将本文提出的算法应用于人脸跟踪任务中,并与文献[6-7]中的基于 CamShift 改进的人脸跟踪算法进行对比实验,实验结果表明本文提出的算法在人脸跟踪任务上有着更好的表现。

1 传统的 CamShift 算法

连续自适应均值漂移算法(continuously adaptive mean-SHIFT, CamShift)算法是一种基于密度函数梯度的非参数估计方法,是基于 MeanShift 算法的改进算法,传统的 CamShift 算法通过图像的颜色特征建立颜色直方图从而获得颜色概率分布图,并通过颜色概率分布图实现运动目标的跟踪, CamShift 算法的具体实现步骤如下:

1) 读取视频序列,将视频帧从 RGB 颜色空间转化到 HSV 颜色空间,并提取该视频帧的 H 分量;

2) 选取需要跟踪的目标,将该目标从 RGB 颜色空间转化到 HSV 颜色空间,提取 H 分量并计算 H 分量的直方图,然后计算直方图在该视频帧的概率分布,从而得到颜色概率分布图,即反向投影图;

3) 在反向投影图上使用 MeanShift 算法进行搜索,设搜索窗口的中心坐标为 (x, y) ,分别计算搜索窗口内像素点的零阶矩和一阶矩,从而获得搜索目标的质心坐标。零阶矩、一阶矩、二阶矩和质心坐标的计算公式分别如式(1)~(4)所示。

零阶矩:

$$M_{00} = \sum_x \sum_y I(x, y) \quad (1)$$

一阶矩:

$$\begin{cases} M_{10} = \sum_x \sum_y xI(x, y) \\ M_{01} = \sum_x \sum_y yI(x, y) \end{cases} \quad (2)$$

二阶矩:

$$\begin{cases} M_{20} = \sum_x \sum_y x^2 I(x, y) \\ M_{02} = \sum_x \sum_y y^2 I(x, y) \\ M_{11} = \sum_x \sum_y xyI(x, y) \end{cases} \quad (3)$$

目标质心坐标:

$$x_d = \frac{M_{01}}{M_{00}}, y_d = \frac{M_{10}}{M_{00}} \quad (4)$$

4) 将搜索窗口的中心移动到质心坐标上,并按照式(5)对搜索窗口的宽度 w 、长度 l 进行调整,式(5)中的 a 、 b 、 c 分别通过式(6)计算得出。

$$\begin{aligned} \omega &= \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}}, \\ l &= \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \quad (5) \\ a &= \frac{M_{20}}{M_{00}} - x_d^2, b = 2 \times \left(\frac{M_{11}}{M_{00}} - x_d y_d \right), c = \frac{M_{02}}{M_{00}} - y_d^2 \quad (6) \end{aligned}$$

5)判断搜索窗口的中心坐标是否收敛(搜索窗中心移动距离小于预设阈值或迭代次数达到设定值),若收敛则返回搜索结果,并将当前的搜索结果作为下一帧初始的搜索窗口,否则回到步骤 3),直到收敛为止。

2 改进的 CamShift 算法

2.1 颜色直方图

传统的 CamShift 将图像从 RGB 颜色空间转换到 HSV 颜色空间,并提取 H 分量根据选定的区域创建色度分量的颜色直方图。文献[10]利用了图像的 H 和 S 分量对目标建立颜色直方图,本文按照文献[10]的方法分别提取图像的色度、饱和度分量即 H 分量、 S 分量,并按照文献[11]的方式划分为 16 级,每级共 16 个色度。根据 MeanShift 算法的原理,目标的颜色特征直方图分布为:

$$\hat{P}_u(y) = \begin{cases} C_h \sum_{i=1}^{n_h} k \left(\left\| \frac{y-x_i}{h} \right\| \right)^2 \delta[b_1(x_i) - u], 1 \leq u \leq 16 \\ C_s \sum_{i=1}^{n_s} k \left(\left\| \frac{y-x_i}{h} \right\| \right)^2 \delta[b_2(x_i) - u], 17 \leq u \leq 32 \end{cases} \quad (7)$$

式中: n_h 、 n_s 分别表示 H 、 S 的分级数量; C_h 、 C_s 为归一化常数,分别可从 $\sum_{i=1}^{n_h} \hat{P}_u(y) = 1$ 、 $\sum_{i=1}^{n_s} \hat{P}_u(y) = 1$ 推出; $b_1(x_i)$ 、 $b_2(x_i)$ 分别表示像素 x_i 的 H 、 S 分量的特征值; $\delta[b_1(x_i) - u]$ 、 $\delta[b_2(x_i) - u]$ 为单位脉冲函数。

2.2 边缘梯度直方图

当目标的颜色和背景的颜色十分相近时,仅使用颜色特征表示跟踪目标往往很难达到理想的跟踪效果。图像的边缘特征能够很好地描述图像的轮廓与纹理,其受光照以及颜色的影响较小。因此将跟踪目标的颜色特征和边缘特征相结合能够提高跟踪的精确度。本文将图像的边缘梯度直方图按 $\pi/8$ 的角度进行分级,共分为 16 级。要得到边缘梯度直方图的首要需要对图像进行灰度变换,得到灰度图像。对灰度图像使用 Scharr 算子对灰度图像进行边缘检测,按照式(8)可以得到像素点 (x, y) 的水平梯度 G_x 、垂直梯度 G_y 、梯度幅值 $G(x, y)$ 和梯度方向 $\theta(x, y)$;

$$\begin{cases} G_x = I(x+1, y) - I(x-1, y) \\ G_y = I(x, y+1) - I(x, y-1) \\ G(x, y) = \sqrt{(G_x)^2 + (G_y)^2} \\ \theta(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \end{cases} \quad (8)$$

目标的边缘梯度特征直方图分布^[16]如式(9)所示,式中 n_e 表示梯度的分级数量; C_e 为归一化常数,可从 $\sum_{i=1}^{n_e} \hat{P}_u(y) = 1$ 推出; $b_3(x_i)$ 表示像素 x_i 的梯度方向的特征值; $\delta[b_3(x_i) - u]$ 为单位脉冲函数。

$$\hat{P}_u(y) = C_e \sum_{i=1}^{n_e} k \left(\left\| \frac{y-x_i}{h} \right\| \right)^2 \delta[b_3(x_i) - u], 33 \leq u \leq 48 \quad (9)$$

将目标的颜色直方图和边缘梯度直方图进行融合,可得到目标的融合直方图分布,如式(10)所示。融合后的直方图能够更好地表示图像的颜色与边缘特征。

$$\hat{P}_u(y) = \begin{cases} C_h \sum_{i=1}^{n_h} k \left(\left\| \frac{y-x_i}{h} \right\| \right)^2 \delta[b_1(x_i) - u], 1 \leq u \leq 16 \\ C_s \sum_{i=1}^{n_s} k \left(\left\| \frac{y-x_i}{h} \right\| \right)^2 \delta[b_2(x_i) - u], 17 \leq u \leq 32 \\ C_e \sum_{i=1}^{n_e} k \left(\left\| \frac{y-x_i}{h} \right\| \right)^2 \delta[b_3(x_i) - u], 33 \leq u \leq 48 \end{cases} \quad (10)$$

2.3 感知哈希

图像哈希算法具有良好的鲁棒性和实时性,在内容搜索、图片搜索、数字水印、模式识别等领域被广泛应用。感知哈希算法^[17]是哈希算法中最使用的算法之一,通过感知哈希算法生成的哈希序列具有唯一性、安全性、快速性等优势,因此其被广泛应用于图像搜索领域。鉴于感知哈希算法的优点,很多学者将感知哈希算法应用于目标跟踪领域^[18-20]。

感知哈希算法通过使用一组二进制比特串来近似描述图像的主体信息,将图像之间的匹配从像素级匹配转化为二进制比特串匹配,这能大大提升图像匹配的效率和很强的实时性和实用性。对于图像 A 和图像 B,只需要比较其通过感知哈希算法生成的哈希序列 h_A 、 h_B ,即可得出两幅图像之间的相似性。感知哈希匹配函数能够量化地表示两幅图像之间的相似程度,并为判断两幅图像是否相似提供了依据。哈希匹配函数如式(11)所示。

$$Distance = pHashMatch(h_A, h_B) \quad (11)$$

式中: $Distance$ 代表两幅图像之间的汉明距离。汉明距离是信息理论中的一种计算两个比特串距离的算法,通过统计等长比特串对应位置的不同字符的个数以表示两个比特串的距离。通过两组感知哈希序列的汉明距离的大小则可计算出两幅图像之间的相似程度。

本文在感知哈希的基础上进行了改进,首先对图像使用感知哈希算法得到整体图像的感知哈希序列 h_k ,然后以图像的中心点对图像进行分割,共分割出 4 个等面积的子区域并分别对 4 个子区域使用感知哈希算法从而可以得到 4 个哈希序列,即 h_1 、 h_2 、 h_3 、 h_4 。从而生成完整的哈希序列 H_{hash} ,如式(12)所示。通过这种方式生成的哈希序列能提

取出图像中更为丰富的信息,生成更具有鲁棒性的感知哈希特征。

$$H_{\text{hash}} = (h_g, h_1, h_2, h_3, h_4) \quad (12)$$

在对两幅图像进行相似计算时,可以分别得到整体图像哈希序列之间的汉明距离 d_g 以及 4 个子区域哈希序列之间的汉明距离,即 d_1, d_2, d_3, d_4 。在得到了汉明距离之后可以通过式(13)计算得出两幅图像之间的相似度。式中 S_{hash} 表示两幅图像之间的相似度, w_1, w_2 分别表示权重, N 表示哈希序列总共包含的比特位数量。本文取 w_1, w_2 分别为 0.5、0.5,取 $N=64$ 。

$$S_{\text{hash}} = w_1 \times \left(1 - \frac{d_g}{N}\right) + w_2 \times \left(1 - \frac{d_1 + d_2 + d_3 + d_4}{4 \times N}\right) \quad (13)$$

本文在使用融合直方图进行初步搜索后,若搜索结果不准确,则使用改进的感知哈希算法进行重新搜索,这样能够提高跟踪算法的准确性。

2.4 SIFT 算法

SIFT 算法提取到的特征点具有良好的尺度不变性和旋转不变性,因此可以很好对两幅图像进行匹配。鉴于 SIFT 算法的优势,很多学者将 SIFT 算法^[21-22]与目标跟踪算法相融合,提高了跟踪的稳定性。

图像熵是一种图像特征的统计形式,它反映了图像中信息的含量。图像的局部熵越大则表示该点附近所含的信息量越大。文献[23-25]通过实验证明了对图像使用局部熵进行分割将图像分割为高熵部分和低熵部分,并只对高熵部分进行 SIFT 特征提取,能够获取原图像大部分的特征点。通过这种改进能够缩小特征提取的区域,从而提升效率。

本文基于上述文献的研究成果对图像进行信息熵分割,只在图像的高信息熵部分进行 SIFT 特征点的提取,从减少提取特征点的耗时,并将提取到的特征点用于跟踪算法中的目标定位。对目标模板与当前视频帧分别进行信息熵分割并在高熵部分进行 SIFT 特征点的提取,并对提取到的特征点进行匹配,从而获得初始的 CamShift 算法搜索窗口,提升跟踪算法的准确率。

2.5 卡尔曼滤波

卡尔曼滤波(Kalman filtering)是一种利用线性系统状态方程,通过系统输入输出观测数据,对系统状态进行最优估计的算法。Kalman 滤波在测量方差已知的情况下能够从一系列存在测量噪声的数据中,估计动态系统的状态。Kalman 滤波分为预测和修正两个阶段,其中预测阶段包括两部分:状态预测和误差协方差预测。对 Kalman 滤波算法建模如下:

信号模型:

$$\mathbf{X}_k = \mathbf{A}_k \mathbf{X}_{k-1} + \mathbf{B}_k \mathbf{W}_k \quad (14)$$

观测模型:

$$\mathbf{Z}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{B}_k \mathbf{V}_k \quad (15)$$

式中: $\mathbf{A}_k, \mathbf{B}_k, \mathbf{H}_k$ 分别是状态转移矩阵、输入矩阵和观测矩阵, $\mathbf{W}_k, \mathbf{V}_k$ 分别为动态噪声和观测噪声; \mathbf{W}_k 与 \mathbf{V}_k 是互不相关的白噪声序列; \mathbf{X}_{k-1} 为状态向量,其中核心部分是 \mathbf{Z}_k , 决定 Kalman 滤波预测的准确性。

在使用 SIFT 算法进行初始搜索框定位时,若跟踪目标外观发生巨大变化时,可能会产生搜索框定位失败现象。因此本文在 SIFT 算法定位失败时采用 Kalman 滤波预测的搜索框作为初始搜索框进行搜索。

3 算法实现

本文在传统的 CamShift 跟踪算法的基础上进行了改进。本文将目标的色调特征、饱和度特征以及梯度特征相结合从而获取目标的融合直方图使得目标直方图的信息更为丰富;同时使用感知哈希在目标外观发生巨大变化时导致直方图信息不够准确的情况下重新进行目标的搜索,本文对感知哈希进行了进一步的改进使得提取到感知哈希特征更具有鲁棒性从而使得跟踪算法更为准确;在对图像进行 SIFT 算法提取特征点前先对图像进行信息熵分割,在图像的高熵部分进行特征点的提取能够提取原始图像大部分特征点的同时加速特征点提取的过程,从而提升特征点提取的效率;在 SIFT 算法完成特征点的提取后对跟踪目标和视频序列的特征点进行匹配从而获得跟踪算法的初始搜索框,进一步提升了跟踪算法的准确率。同时引入卡尔曼滤波预测目标的运动,当匹配的特征点数量过时,则使用卡尔曼滤波预测的搜索框作为初始的搜索框,从而使得跟踪算法更为准确。

本文算法的具体步骤如下:

- 1) 读取视频帧,将图像从 RGB 颜色空间转换到 HSV 颜色空间;
- 2) 选取目标,按照式(10)得到目标的融合直方图;
- 3) 读取下一帧图像,并根据初始搜索窗口在 CamShift 跟踪算法的框架下进行搜索,获取候选目标的大小和位置;
- 4) 获取候选目标的融合直方图,通过计算可以得到候选目标与目标模板之间的 Bhattacharyya 距离。若 Bhattacharyya 小于阈值则将其作为跟踪结果并进入步骤 6); 否则认为 CamShift 跟踪算法获取的候选目标不准确并进入步骤 5);
- 5) 使用改进的感知哈希算法提取目标模板的感知哈希序列,在 CamShift 跟踪算法得到的位置附近进行搜索,选取与目标模板感知哈希序列之间汉明距离最小的候选目标,并将该候选目标作为跟踪结果并进入步骤 6);
- 6) 用 SIFT 算法在图像的高熵部分提取跟踪目标与下一帧的特征点并进行匹配,得到特征点的最小外接矩形,将其作为下一帧的初始搜索窗口进行搜索并进入步骤 8)。若 SIFT 算法匹配失败导致无法获取最小外接矩形则进入步骤 7);
- 7) 使用 Kalman 滤波预测目标可能出现和将其作为下一帧序列中的初始搜索窗口进行搜索并进入步骤 8);

8)重复步骤 3)~7)直到视频序列结束。

4 实验结果与分析

4.1 改进的感知哈希实验分析

本文对改进后的感知哈希进行了实验并与感知哈希算法进行对比,测试了改进后的感知哈希在不同场景的性能。实验使用的图片如图 1 所示,实验数据如表 1 所示。从表 1 的实验数据可知改进后的感知哈希比起感知哈希提取到了更丰富的感知哈希特征,从而使得目标在受到外界干扰以及发生形变或尺度变化的情况下,目标变化前后图片之间相似度得到了明显得提升。因此将改进后的感知哈希算法应用到目标跟踪算法中,能够提高当目标视觉外观发生改变时目标变化前后的相似度,提升搜索目标时算法的正确性,从而提升跟踪算法的准确性。

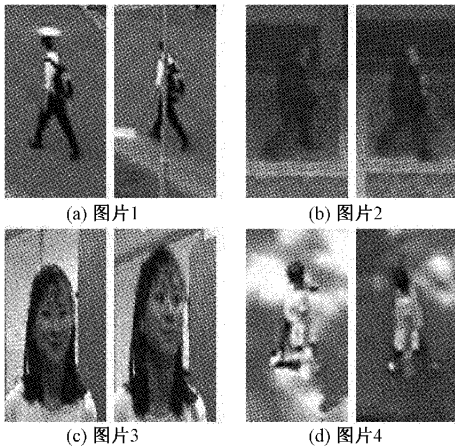


图 1 目标变化前后对比图

表 1 目标变化前后相似度

实验	干扰因素	感知哈/%	改进的感知哈希/%
a	遮挡	71.8	88.8
b	形变	54.6	73.2
c	尺度变化	70.3	80.4
d	光照变化	60.9	75.3

4.2 跟踪算法评价

本文使用文献[26]中的评价指标来评估本文算法的性能,所采取的评价指标为成功率和距离精度。成功率是指预测框(记作 A)和真实框(记作 B)的重叠率(overlap precision, OP)大于设定阈值的视频帧数占整个视频序列帧数的比例。重叠精度可通过计算预测框和真实框之间的交并比得出,其计算公式如式(16)所示。距离精度是指中心位置误差(center location error, CLE)小于设定阈值的视频帧数占整个视频序列帧数的比例。中心位置误差可通过计算预测目标的中心坐标(x_p, y_p)和真实目标的中心坐标(x_g, y_g)之间的欧式距离得到,其计算公式如式(17)所示。

$$OP = \frac{|A \cap B|}{|A \cup B|} \tag{16}$$

$$CLE = \sqrt{(x_g - x_p)^2 + (y_g - y_p)^2} \tag{17}$$

本文在 OTB-100 视频序列中选取了 8 个视频序列与常用的跟踪算法进行对比实验。测试视频序列如表 2 所示。

表 2 所使用的视频序列描述

视频序列	序列帧数	图像尺寸	干扰因素
Bolt	350	640×360	光照变化、相似背景
CarDark	393	320×240	光照变化、相似背景
Crowds	347	600×480	光照变化、形变、相似背景
Girl	500	128×96	遮挡、旋转、尺度变化
KiteSurf	84	480×270	光照变化、旋转、遮挡
Singer2	366	624×352	光照变化、相似背景、旋转
Subway	176	352×288	遮挡、相似背景
Walking	412	768×576	遮挡、尺度变化

4.3 定性分析

1)实验 1 针对跟踪目标在受到遮挡的情况下测试本文算法的跟踪准确率。选取的实验视频序列为 Girl、SubWay、Walking、KiteSurf,其中视频序列 Girl 的实验效果图如图 2 所示。在视频序列 Girl 中跟踪的目标对象为女孩,从图 2 中可以发现 KCF、TLD、传统的 CamShift 跟踪算法都产生了跟踪不准确的现象。其中 KCF 算法在真实目标经过遮挡后错误地将另一个相似目标作为跟踪目标,丢失了真实目标,CamShift 跟踪算法和 TLD 算法在跟踪目标时发生了漂移现象。从图 2 中可以看出本文提出的算法能够准确地遮挡情况下稳定地跟踪目标。

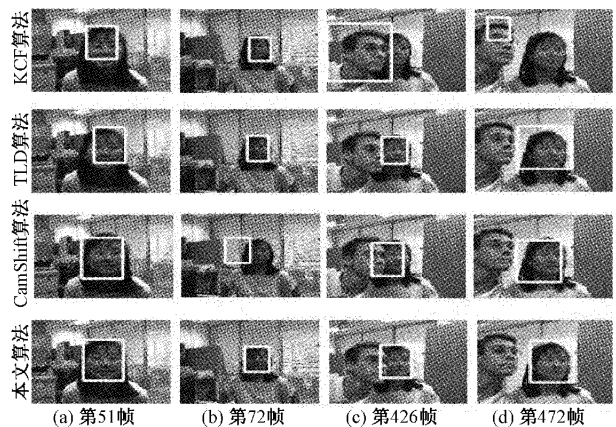


图 2 视频序列 Girl 跟踪结果

2)实验 2 针对跟踪目标在受到相似背景干扰的情况下测试本文算法的跟踪准确率。选取的实验视频序列为 Bolt、CarDark,其中视频序列 Bolt 的实验效果图如图 3 所示。在视频序列 Bolt 中跟踪的目标对象为身穿黄色运动

服的博尔特,在该视频序列中存在另一位身穿黄色运动服的运动员,对目标进行跟踪时与目标相似的背景会产生一定的干扰作用。从图3中可以发现传统的CamShift跟踪算法在第237和330帧均产生较大的跟踪漂移现象;KCF算法在第26帧丢失了跟踪目标;TLD算法在第237帧时产生了跟踪不准确的现象。本文提出的算法在跟踪时受到相似背景的影响较小,在跟踪过程中能够始终稳定准确地跟踪目标,在相似目标的干扰下有着良好的性能。

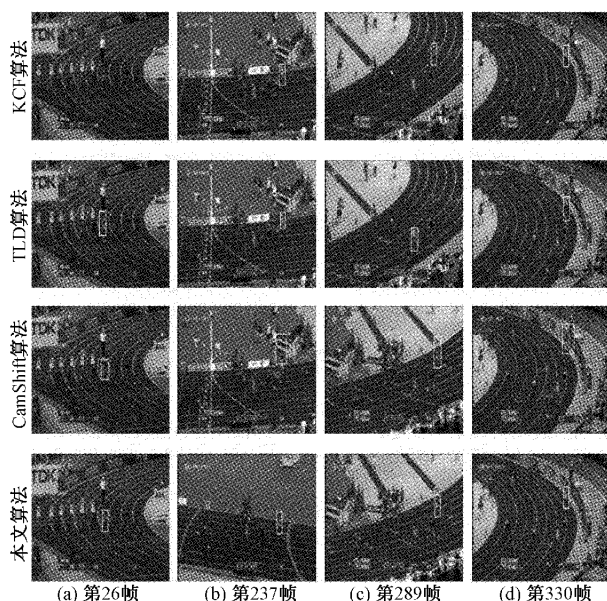


图3 视频序列 Bolt 跟踪结果

3)实验3针对跟踪目标在受到光照变化干扰的情况下测试本文算法的跟踪准确率。选取的实验视频序列为Crowds、Singer2,其中视频序列Singer2的实验效果如图4所示。在视频序列Singer2中跟踪的目标对象为身穿黑色西装的歌手,从图4中可以看出跟踪目标在第58、59帧和第329、330帧经历剧烈的光照变化。KCF算法在第58、329帧产生跟踪不准确的现象,在第59、330帧丢失了跟踪目标;TLD算法在第59、330帧因为光照的剧烈变化同样丢失了跟踪目标;传统的CamShift跟踪算法则在跟踪过程中表现得不稳定,在第329、330帧中产生了较大的偏移。本文提出的跟踪算法相在跟踪过程中较为稳定,受到光照剧烈变化的影响较小,能够准确地跟踪目标,在光照剧烈变化的影响下有着良好的跟踪准确度。

4.4 定量分析

为了定量分析本文算法的性能,将本文算法和其他跟踪算法在选取的8个视频序列上进行对比实验,跟踪算法的成功率和距离精度分别如表3和4所示,其中重叠率阈值取0.5,中心坐标误差阈值取20。从两张表的实验数据中可以发现本文算法在跟踪成功率和距离精度上与其他算法相比有着显著的提升。

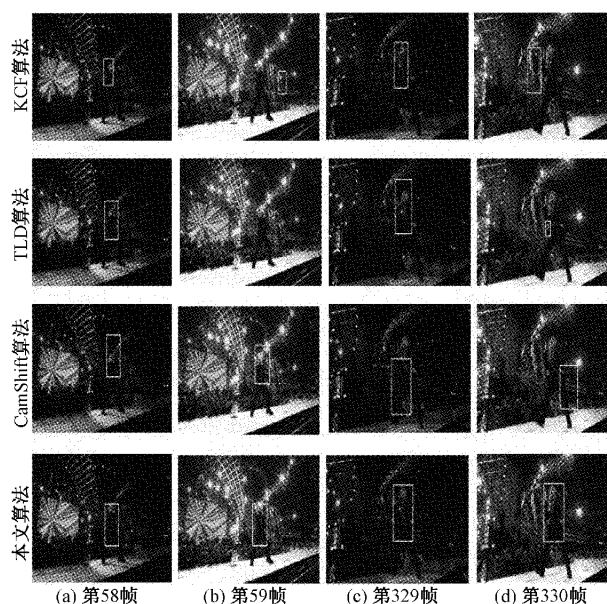


图4 视频序列 Singer2 跟踪结果

表3 各种跟踪算法成功率对比 %

视频序列	TLD	KCF	CamShift	本文算法
Bolt	64.10	48.90	32.90	92.00
CarDark	90.50	82.30	26.70	93.10
Crowds	65.60	62.10	25.40	88.10
Girl	71.70	58.30	47.20	90.80
KiteSurf	42.10	48.10	26.50	88.00
Singer2	57.20	67.10	36.10	90.70
Subway	70.60	51.10	29.30	90.80
Walking	59.80	54.50	36.70	87.70
平均值	65.20	59.05	32.60	90.15

表4 各种跟踪算法距离精度对比 %

视频序列	TLD	KCF	CamShift	本文算法
Bolt	76.50	56.10	61.30	96.20
CarDark	99.70	93.10	52.80	96.10
Crowds	94.20	86.10	50.00	96.20
Girl	97.50	87.90	77.30	93.60
KiteSurf	90.30	85.50	50.60	94.00
Singer2	50.90	58.60	33.40	90.70
Subway	92.50	90.20	44.20	94.80
Walking	87.80	89.00	47.40	94.60
平均值	86.18	80.81	52.13	94.53

4.5 综合分析

为了进一步分析本文跟踪算法的性能,在8个视频序列上分别采用不同的重叠率阈值和中心坐标误差阈值和其他跟踪算法进行对比实验,成功率曲线图和距离精度曲线

图分别如图 5 和 6 所示。从曲线图可以本文跟踪算法有着良好的跟踪准确率和稳定性。

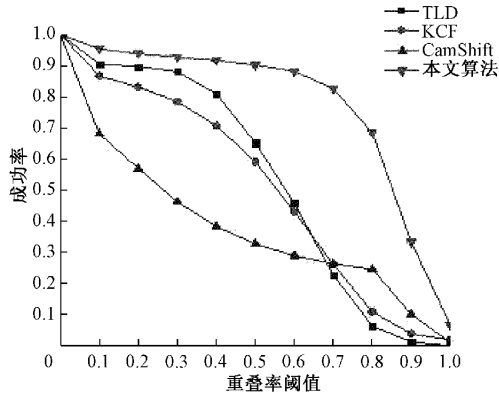


图 5 成功率曲线图

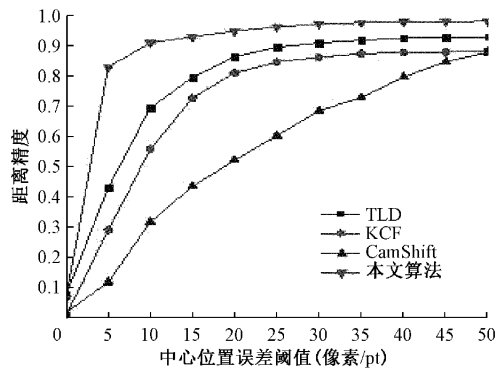


图 6 距离精度曲线图

4.6 应用分析

在对本文提出的跟踪算法的性能进行测试之后,进一步将跟踪算法应用于人脸跟踪任务中,并与文献[6-7]中的人脸跟踪算法进行对比实验。本文选取 OTB-100 数据集中 10 个跟踪目标为人脸的视频序列作为测试数据,即 BlurFace、David、David2、Dudek、Jumping、Man、Mhyang、Trellis、Boy、FleetFace。文献[6]中在 CamShift 算法的基础上引入了卡尔曼滤波对人脸的运动进行了估计一定程度上提升了准确率,然而当仅使用图像的直方图进行人脸跟踪,当目标的外观模型发生巨大变化时直方图的信息和当前目标的信息之间的差异同样会增大,因此这可能导致跟踪不准确。文献[7]中将目标的纹理信息和卡尔曼滤波与 CamShift 算法相结合,一定程度上丰富了目标的特征信息,然而目标的纹理信息属于图像的高频信息即图像的细节部分,当光照强度过大或过小以及运动模糊时可能导致目标的细节信息严重丢失,从而导致跟踪精度下降。本文将改进的感知哈希和 CamShift 算法相结合,改进的感知哈希能够更好地表达目标的主体特征信息。并使用 SIFT 算法对特征点进行匹配,提升了人脸跟踪时的准确率。在重叠率阈值取 0.5,中心坐标误差阈值取 20 的情况下,成功

率和距离精度分别达到了 93.5% 和 96.2%。其中视频序列 Trellis 和 Dudek 的实验效果对比分别如图 7 和 8 所示。



图 7 视频序列 Trellis 跟踪结果

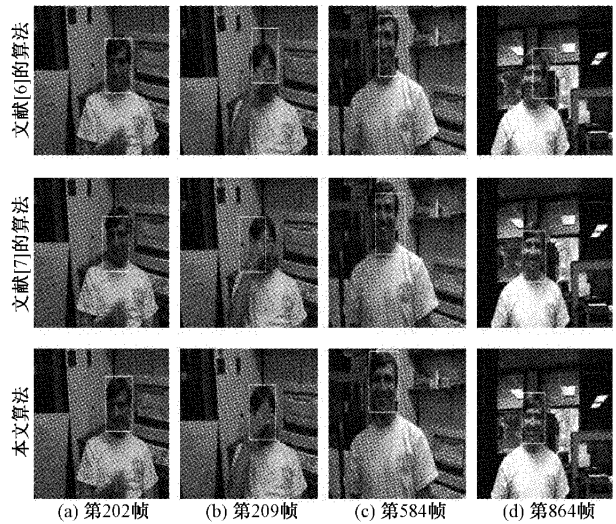


图 8 视频序列 Dudek 跟踪结果

在视频序列 Trellis 中人脸部的光照强度信息发生了多次改变,先从弱到强再由强到弱,文献[6-7]的人脸跟踪算法在受到光照干扰时会受到一定地干扰从而导致跟踪框的偏移,本文的跟踪算法在应用于人脸跟踪时受到光照变化的影响相对较小,能够稳定地对人脸进行跟踪。

在视频序列 Dudek 中人脸部的表情发生了改变,同时,在第 209 帧中手对脸部进行了遮挡。在第 584 和 864 帧人所处的场景发生了改变。以上因素一定程度上会干扰跟踪结果。文献[6-7]的人脸跟踪算法在受到手部遮挡时均产生了偏移现象,在场景改变时也产生了跟踪不准确的现象。本文的跟踪算法在应用于人脸跟踪时受到遮挡以及场景变化的影响相对较小,能够准确地跟踪人脸目标。

本文算法和文献[6-7]中的人脸跟踪算法在选取的 10 个人脸视频序列的成功率曲线图和距离曲线图分别如图 9 和 10 所示。从曲线图和实验效果图中可以看出本文的算法在人脸跟踪任务中有着良好的性能,能够准确稳定地对人脸目标进行跟踪。

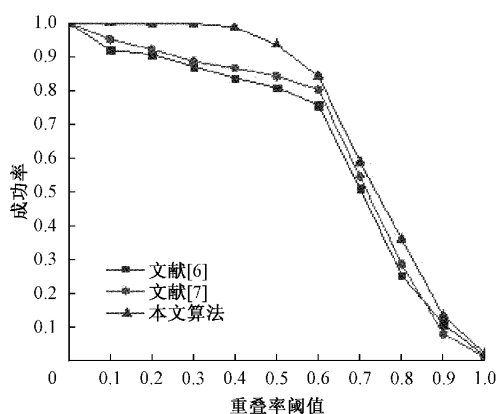


图 9 人脸跟踪成功率曲线图

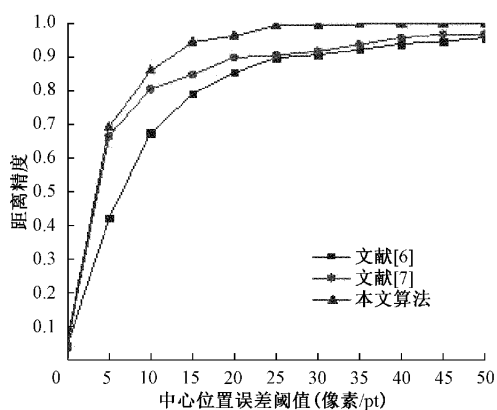


图 10 人脸跟踪距离精度曲线图

5 结 论

本文针对传统的 CamShift 仅使用目标的颜色直方图作为特征,在相似背景干扰、遮挡、高速运动等情况下容易出现跟踪不准确或丢失跟踪目标现象的不足,提出了基于 SIFT 和感知哈希改进的 CamShift 跟踪算法。首先,将目标从 RGB 颜色空间转为 HSV 颜色模空间,分别得出色调和饱和度直方图并提取图像的边缘梯度直方图进行融合获得目标的融合直方图。然后在 CamShift 算法框架下得到最优候选目标,若候选目标与目标模板的 Bhattacharyya 距离小于阈值时,则使用改进的感知哈希进行最优候选目标的搜索。在下一帧搜索时,使用 SIFT 算法进行匹配从而获取初始搜索窗口,若 SIFT 算法匹配失败,则使用卡尔曼滤波预测的搜索框作为初始搜索窗口。通过和常用的跟踪算法进行对比实验,实验表明本文提出的算法能够在遮挡、光照变化以及相似背景的干扰下稳定地跟踪目标,具有很

高的准确性。在对本文提出的跟踪算法的性能进行测试之后,将本文提出的算法进一步应用于人脸跟踪任务中,并和其他的人脸跟踪算法进行对比实验,实验结果表明本文提出的算法相比其他算法有着更好的表现,具有一定的实际意义和应用价值。

参考文献

- [1] 李晓明,黄慧,应毅,等.基于图像置信度修正的智能小车导航目标快速跟踪算法[J].国外电子测量技术,2022,41(2):162-168.
- [2] 张雨,陈熙源,朱敏,等.基于斜率约束和回溯搜索的水下多目标跟踪方法[J].仪器仪表学报,2021,42(9):81-88.
- [3] 李建良,张婷婷,陶知非,等.基于改进 Camshift 与 Kalman 滤波融合的领航车辆跟踪算法[J].电子测量与仪器学报,2021,35(6):131-139.
- [4] BANKAR R, SALANKAR S. Improvement of head gesture recognition using camshift based face tracking with UKF[C]. 2019 9th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-19), IEEE, 2019: 1-5.
- [5] 王俊南,赖惠成,杨东旭.基于肤色分割的改进 Camshift 人脸跟踪算法[J].现代电子技术,2019,42(10):166-170.
- [6] 田小锋,董佳佳,封硕.基于卡尔曼滤波运动估计的 Camshift 人脸跟踪方法[J].信息技术与信息化,2021(8):99-101.
- [7] 张娜,王瑜,孙晓荣.融合纹理信息和预测机制的人脸跟踪算法[J].计算机工程与设计,2017,38(7):1934-1937.
- [8] ATOEV S, NISHANOV A, ANDIRAZAKOV F. Object tracking method based on Kalman filter and camshift algorithm for UAV applications[C]. 2021 International Conference on Information Science and Communications Technologies(ICISCT), IEEE, 2021: 1-4.
- [9] 沈思源,李震霄,孙伟.基于改进 Camshift 的无人机目标跟踪算法研究[J].微电子学与计算机,2019,36(11):76-83.
- [10] DU S, XU H, LI T. Implementation of camshift target tracking algorithm based on hybrid filtering and multifeature fusion[J]. Journal of Sensors, 2020: 1-13.
- [11] 李凌.融合颜色和边缘特征的改进型 CamShift 跟踪算法[J].激光杂志,2017,38(7):138-142.
- [12] 吴玮,郑娟毅,杜乐.多特征融合的 Camshift 运动目标跟踪算法[J].计算机科学,2018,45(7):252-258.
- [13] 郭金朋,战荫伟,李雪武.基于 SURF 和 Camshift 的运

- 动目标跟踪[J]. 计算机工程与设计, 2017, 38(8): 2190-2193.
- [14] 李杰超, 张潇宵, 王凯. 改进局部二值模式算法与 Camshift 结合的目标跟踪方法[J]. 科学技术与工程, 2021, 21(26): 11232-11239.
- [15] 于国庆, 吴赛敏. 结合 Kalman 滤波和 LBP 纹理的 CAMShift 改进算法[J]. 现代电子技术, 2020, 43(12): 65-68.
- [16] 马正华, 顾苏杭, 戎海龙. 基于 SIFT 特征匹配的 CamShift 运动目标跟踪算法[J]. 计算机科学, 2014, 41(6): 291-294.
- [17] SAMANTA P, JAIN S. Analysis of perceptual hashing algorithms in image manipulation detection [J]. Procedia Computer Science, 2021, 185(2): 203-212.
- [18] 陈优良, 肖钢, 卞焕, 等. 一种融合动态预测的感知哈希目标跟踪算法[J]. 测绘通报, 2020(2): 17-23.
- [19] 李子印, 朱明凌, 陈柱. 融合图像感知哈希技术的运动目标跟踪[J]. 中国图象图形学报, 2015, 20(6): 795-804.
- [20] 张立国, 王松, 金梅, 等. 基于多尺度感知哈希特征的目标跟踪算法研究[J]. 高技术通讯, 2018, 28(3): 219-226.
- [21] BISWAS B, KR GHOSH S, HORE M, et al. Sift-based visual tracking using optical flow and belief propagation algorithm [J]. The Computer Journal, 2022, 65(1): 1-17.
- [22] 王智军, 王建华. 基于 SIFT 验证的 Mean Shift 跟踪运动目标新算法[J]. 电光与控制, 2016, 23(11): 93-96.
- [23] 刘雪亭, 李太君, 肖沙, 等. 一种融合信息熵和改进 SIFT 算法的图像检索方法[J]. 河南科技大学学报(自然科学版), 2014, 35(6): 42-46, 7.
- [24] 冀振宇. 基于局部统计特性及 SIFT 算法的图像配准[J]. 太原理工大学学报, 2017, 48(5): 828-832.
- [25] 刘自金, 石玉英. 基于信息熵的改进 SIFT 图像快速匹配算法[J]. 激光杂志, 2021, 42(12): 129-135.
- [26] WU Y, LIM J, YANG M H. Online object tracking: A benchmark[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2411-2418.

作者简介

李文举, 博士, 教授, 主要研究方向为计算机视觉、模式识别、智能检测。

E-mail: wjli@sit.edu.cn

王子杰, 硕士研究生, 主要研究方向为计算机视觉。

E-mail: 704021267@qq.com

崔柳(通信作者), 博士, 讲师, 主要研究方向为制导导航与控制, 微传感器和生物医学信号处理、图像处理。

E-mail: cui Liu8475@msn.com