

基于多尺度感知哈希特征的目标跟踪算法研究^①

张立国^② 王 松^③ 金 梅 崔利洋

(燕山大学河北省测试计量技术与仪器重点实验室 秦皇岛 066004)

摘 要 研究了目标跟踪与检测。考虑到其应用环境复杂,造成对光照变化鲁棒性不强,目标跟踪过程中数据丢失,实时跟踪时检测速度不够的问题,结合哈希指纹特征对光照变化鲁棒性强的特点,提出了一种基于多尺度(multi-scale)感知哈希(Hash)特征的目标跟踪算法,即 MHash 算法。该算法采用多种尺度扫描窗口对待检测图像进行扫描,利用双线性内插值的方法缩小扫描窗口提取哈希指纹特征;同时利用汉明距离与高斯分布相结合的方法对扫描窗口进行评价,评分最高的作为目标窗口。该算法已应用于人脸、运动车辆及快速和缓慢移动物体的跟踪与检测,并分析了跟踪与检测结果。结果表明,MHash 算法具有速度快、跟踪稳定及对光照、目标尺度变化鲁棒性强的特点,能够长期有效地对目标进行跟踪。

关键词 目标跟踪,多尺度,感知哈希,汉明距离,特征提取

0 引 言

目标跟踪一直是图像处理与计算机视觉领域的研究热点,在运动分析^[1]、行为识别^[2,3]、监视^[4]和人机交互^[5]等方面吸引了大量国内外学者投入研究^[6-8]。目前目标跟踪算法大致可以分为两大类:基于生成模型的目标跟踪和基于判别模型的目标跟踪^[9]。基于生成模型的方法主要有均值漂移(meanshift)法^[10]、粒子滤波(particle filter)法^[11]等。判别模型方法主要包括实时压缩跟踪(compressive tracking,CT)^[12]算法、跟踪学习检测(tracking learning detection,TLD)^[13]算法和深度学习(deep learning)类方法。尤其随着深度学习的发展,基于深度学习的目标跟踪算法成为目标跟踪系统未来的发展趋势^[14]。但是,基于深度学习的目标跟踪算法,就其实用程度而言,当前还不能同时满足鲁棒性、实时性和精准度的需要^[15]。其他目标跟踪算法也面临

一系列的挑战,例如遮挡、相似背景、光照变化和姿态变化^[16]等因素导致目标漂移甚至跟踪失败。近年来随着机器视觉技术的发展,基于特征的目标跟踪算法^[17],由于其通过提取图像特征的方法对目标进行识别跟踪,有效地改善了上述问题。

感知哈希特征^[18]提取方法能够很好地描述图像,在相似图像搜索领域有突出表现^[19],因此得到了广泛应用。基于感知哈希特征提取方法的优点,Yu^[20]等人首次将其与粒子滤波相结合应用在目标跟踪中。基于图像、声音等信号在离散余弦变换后的能量都集中在低频部分的特性,提取变换后图像的低频部分,并计算得到哈希指纹特征,通过粒子滤波获取待检测图像的目标位置。该算法的优点是对于目标的光照变化、姿态变化具有较好的鲁棒性,缺点是由于扫描窗口的限制和采用频率图像的低频感兴趣区域(ROI)部分的原因,不能适应目标的尺度变化。另外,由于对图像进行离散余弦变换计算量较大,所以跟踪目标的实时性较差。

① 燕山大学基础研究专项课题(16LGA007)资助项目。
② 男,1978年生,博士,副教授,硕士生导师;研究方向:机器视觉,智能机器人,虚拟现实等;E-mail: 15033572808@163.com
③ 通信作者,E-mail: wszxtf@126.com
(收稿日期:2017-09-25)

基于上述方法中存在的计算速度慢、不能适应目标的尺度变化等问题,同时结合哈希指纹特征具有对光照变化鲁棒性强的优点,本文提出了一种基于多尺度(multi scale)感知哈希(Hash)特征的目标跟踪算法,简称为 MHash 算法。该算法采用双线性内插值的方法缩小扫描窗口,并计算其哈希指纹特征,避免了对图像进行离散余弦变换,因此有效地提高了算法的实时性。而且通过建立多种尺度的扫描窗口,获得多尺度感知哈希指纹特征,增加了算法对目标尺度变化的鲁棒性。本文也基于多尺度哈希指纹特征,提出了一种利用高斯分布与汉明距离相结合的目标评价方法,提高了目标跟踪的准确性。

1 多尺度感知哈希特征提取

本文提出的多尺度感知哈希特征的目标跟踪算法采用多尺度扫描窗口的策略对整幅图像进行扫描。首先,提取某一尺度下的哈希指纹特征组,实现流程如图 1 所示。 $t+1$ 帧图像为待检测图像,方框为某一尺度下遍历整幅图像的扫描窗口,通过双线性内插值的方法将此扫描窗口提取成 8×8 的哈希指纹特征,将整幅图像的哈希指纹特征组合到一起得到该尺度下的哈希指纹特征组。其次,将多种尺度下的哈希指纹特征组合到一起,构成 $t+1$ 帧图像的多尺度哈希指纹特征组如图 2 所示。最后,将 $t+1$ 帧图像获得的多尺度哈希指纹特征组中的哈希

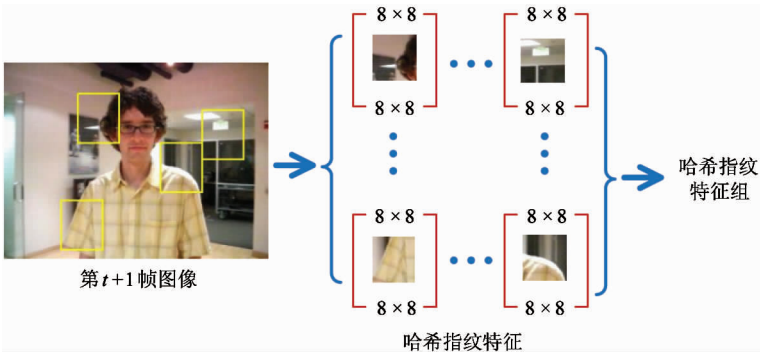


图 1 获取哈希特征组

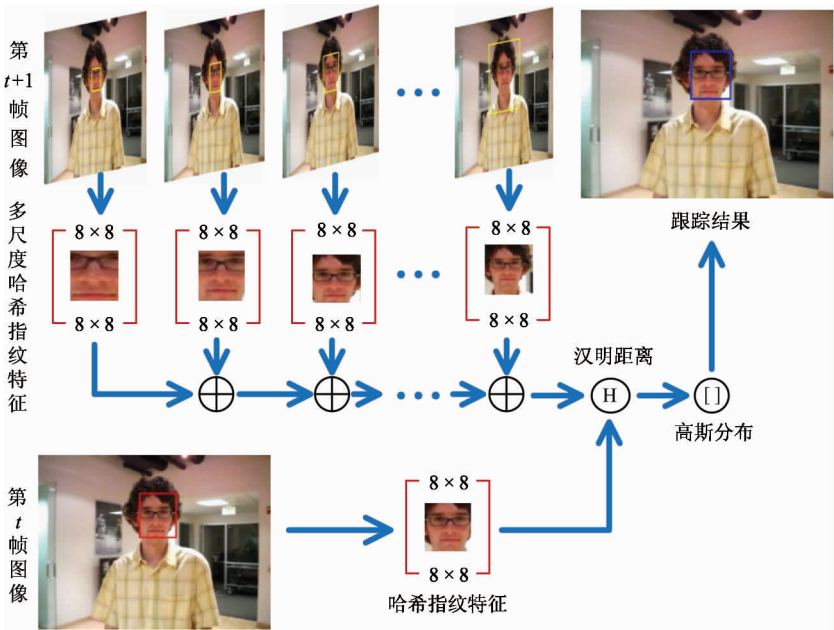


图 2 多尺度图像感知哈希跟踪

指纹特征分别与 t 帧图像目标窗口的哈希指纹特征进行匹配,利用汉明距离和高斯分布相结合的方法对其进行评价,评分最高的窗口作为 $t + 1$ 帧图像的目标窗口,如图 2 中的方框。

1.1 图像预处理

哈希特征是对图像的整体进行描述,忽略细节部分。一张图像的细节是通过高频信息体现的,所以本文的图像预处理就是滤掉图像中的高频部分。高斯低通滤波是对整幅图像进行加权平均,可以很好地保存哈希特征,因此采用高斯低通滤波对输入的图像进行预处理,抑制图像噪声,滤掉图像高频成分。其公式表示如下:

$$F(x, y) = G(x, y, \sigma) * I(x, y) \quad (1)$$

其中 $I(x, y)$ 为待处理的图像, $F(x, y)$ 为经过高斯低通滤波后的图像,高斯函数 $G(x, y, \sigma)$ 表示为

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp \frac{x^2+y^2}{2\sigma^2} \quad (2)$$

其中 σ 是标准差。通过高斯低通滤波处理后的图像滤掉了图像中的高频部分,减少了噪声的干扰,使图像变得更平滑,这有利于后期进行哈希指纹特征的提取。

1.2 多尺度化扫描窗口

对于连续采集的动态图像,目标的尺度变化主要是以上一帧目标窗口为基础进行缩小或放大。因此要提高目标尺度变化的鲁棒性,只需连续地跟随上一帧目标窗口的尺度变化。本文通过对输入目标窗口乘以一个以 1 为基准的系数序列,获得多种尺度的扫描窗口来跟随目标的尺度变化。向上平移一个单位的正切函数曲线由于其在对称点附近具有连续、平缓的特点,可以很好地满足本文对尺度系数取值的要求,尺度系数 S 公式为

$$S = \tan i + 1 \quad (3)$$

其中 i 为变换系数。 i 的取值越密集、范围越大算法对目标尺度变化的鲁棒性越强,但是计算的耗费时间也会越多。算法执行所需要的时间与尺度系数的关系如图 3 所示。本文中综合考虑到目标跟踪的实时性以及目标尺度变化的鲁棒性, i 的取值为以 0.1 为步长在 $[-0.7, 0.7]$ 范围内取 15 个采样点。获得 15 个尺度系数 S 如图 4 所示。

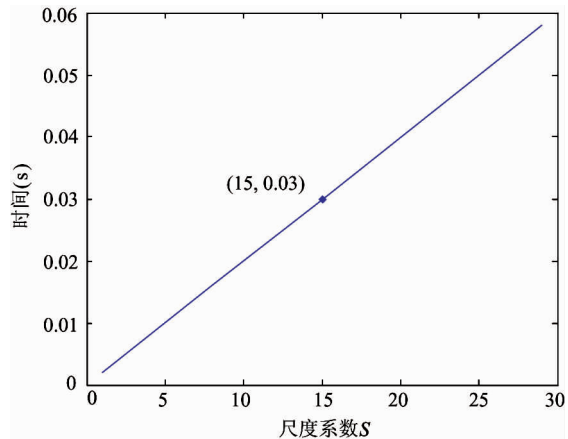


图 3 执行时间与尺度系数关系曲线

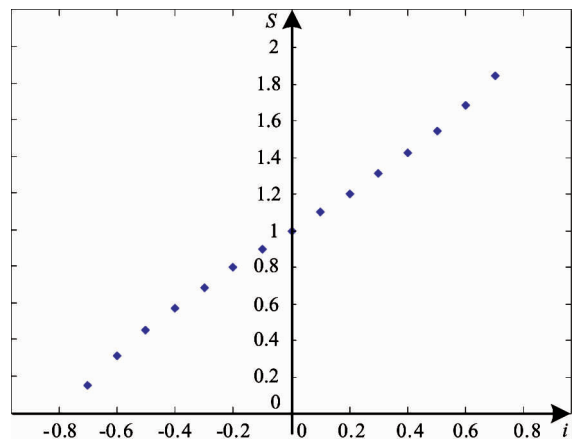


图 4 尺度系数曲线

获取多尺度扫描窗口的公式为

$$SearchWindow = S \times LastWindow \quad (4)$$

其中 $SearchWindow$ 表示尺度变换后的扫描窗口, $LastWindow$ 表示上一帧检测到的目标窗口。 $LastWindow$ 经过尺度化以后获得 15 种不同尺度的扫描窗口。

1.3 双线性内插值哈希指纹特征提取

以上述 15 种不同尺度的扫描窗口对整幅待检测图像进行扫描。将扫描窗口缩小为 8×8 的图像块,即 64 位的特征码。双线性内插值算法是通过获取 4 个最邻近目标像素的像素值再乘以权重系数实现对图像缩放。利用该图像缩放算法提取的 64 位特征码均匀地分布在扫描窗口中,可以更好地描述扫描窗口。同时该方法只需要进行 64 次特征码的转换运算,有效地降低了算法的复杂度,提高了算法的实时性。扫描窗口的第 i_1 行与 64 位特征码的第

i_2 行转换关系式为

$$i_1 + u = i_2 \times \frac{H_1}{H_2} \tag{5}$$

扫描窗口的第 j_1 列与 64 位特征码的第 j_2 列转换关系式为

$$j_1 + v = j_2 \times \frac{W_1}{W_2} \tag{6}$$

其中 i_1 和 j_1 分别为扫描窗口中对应的第 i 行第 j 列, H_1 和 W_1 分别为扫描窗口图像的行数和列数, i_2 和 j_2 分别为缩小后的图像块第 i 行和第 j 列, H_2 和 W_2 分别为缩小后图像块的行数和列数, 在本文中都取 8, u 和 v 分别表示生成的小数部分。64 位特征码中的每一位的灰度值都是与其相邻的 4 个像素的灰度值的平均值, 图像块在 (i, j) 点的灰度值计算公式为

$$I(i_2, j_2) = w_1 S(i_1, j_1) + w_2 S(i_1, j_1 + 1) + w_3 S(i_1 + 1, j_1) + w_4 S(i_1 + 1, j_1 + 1) \tag{7}$$

其中

$$\begin{cases} w_1 = (1 - u)(1 - v) \\ w_2 = (1 - u)v \\ w_3 = u(1 - v) \\ w_4 = uv \end{cases} \tag{8}$$

$S(i_1, j_1)$ 为扫描窗口在第 i 行第 j 列的像素灰度值, $I(i_2, j_2)$ 为缩小后的 64 位特征码在第 i 行第 j 列的像素灰度值, w_1, w_2, w_3, w_4 分别为 $S(i_1, j_1)$ 和其右下角相邻的 3 个像素上的权重。通过上述方法将多尺度的扫描窗口缩小为 8×8 的像素块。

计算获得 8×8 窗口的 64 个特征码的平均值,

遍历该 64 个特征码, 大于平均值的位置置 1, 小于或者等于平均值的位置置 0。这样就将 15 种尺度下的每一个扫描窗口的 64 位特征码转换成了一个用 0 和 1 表示的 64 位哈希指纹特征。将同一尺度下的哈希指纹特征组合到一起得到该图像在该尺度下的哈希指纹特征组。通过提取不同尺度下的哈希指纹特征组获得该图像的多尺度哈希指纹特征组。

2 多尺度哈希指纹特征匹配

单一的利用汉明距离对哈希指纹特征进行评价时存在准确度不够的问题, 本文提出汉明距离与高斯分布相结合的算法, 有效地提高了跟踪的准确度。首先, 以汉明距离作为评价参数提取不同尺度下的最优目标窗口。利用目标窗口的哈希指纹 H_{object} 与待检测图像在不同尺度下的哈希指纹特征 H_{sample}^m 相匹配, 获得目标窗口与待检测图像窗口之间的汉明距离 D_h^m , 将其作为目标匹配的评价参数。汉明距离表示两个 64 位的哈希指纹特征对应位不同的数量。汉明距离的计算公式为

$$D_h^m = \sum_{i=1}^n [H_{\text{object}}(i) \oplus H_{\text{sample}}^m(i)] \tag{9}$$

其中 $m=1, 2, \dots$ 为多尺度哈希指纹特征组的第 m 个哈希指纹特征, D_h^m 表示第 m 个扫描窗口与目标窗口的汉明距离, $H_{\text{object}}(i)$ 表示目标窗口的 64 位哈希指纹特征的第 i 位, $H_{\text{sample}}^m(i)$ 表示第 m 个扫描窗口的 64 位哈希指纹特征的第 i 位, $\oplus$ 表示异或操作。计算某一尺度的汉明距离的流程如图 5 所示。

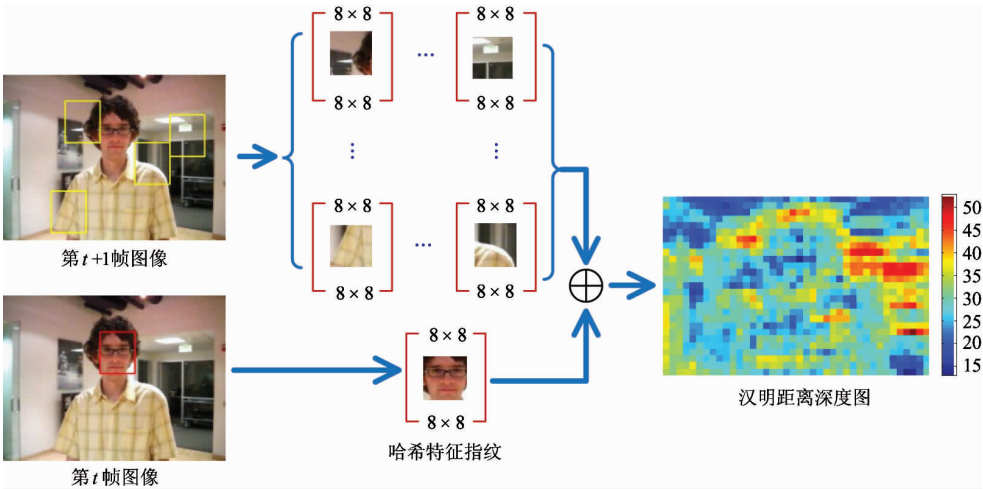


图 5 汉明距离计算流程图

汉明距离^[21]越短说明扫描窗口(采样窗口)与目标窗口的相似度越高。提取每种尺度下汉明距离最短即相似度最高的3个窗口,共获得45个与目标相似度最高的扫描窗口。

然后,通过添加高斯权重的方法,对获得的45个最优扫描窗口进行再评价。对整个待检测图像建立高斯分布模型,以上一帧目标的中心坐标点为高斯模型的中心,最优扫描窗口中心到高斯中心的欧氏距离越远其高斯权重越低。高斯模型的分布公式为

$$g(D_o) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{D_o^2}{2\sigma^2}\right) \quad (10)$$

其中 D_o 表示扫描窗口中心到上一帧目标窗口中心的欧氏距离。 σ 表示最优扫描窗口中心到高斯中心欧式距离的标准差。优化评价的公式为

$$D_{hg} = (64 - D_h^m) \times g(D_o) \quad (11)$$

其中 D_{hg}^m 表示第 m 个扫描窗口高斯分布优化后的评价参数, $64 - D_h^m$ 表示反向汉明距离。反向汉明距离越大说明扫描窗口(采样窗口)与目标窗口的相似度越高。以参数 D_{hg}^m 对45个扫描窗口进行评价, D_{hg}^m 越大表示该扫描窗口的评价得分越高。通过对 D_{hg}^m 排序获得评价得分最高的扫描窗口,作为当前待检测图像中的目标位置窗口。通过高斯模型对扫描窗口的反向汉明距离进行优化前后对比如图6所示。图6(a)表示高斯模型优化前的反向汉明距离深度图,图6(b)表示高斯模型优化后的距离深度图。图中虚线为45个扫描窗口中心在图像中的对应区域。通过对比分析图6(a)和图6(b)可以看出经过高斯模型优化后增加了目标窗口的权值,降低了干扰窗口的权值,有效地滤掉了其他扫描窗口的干扰,突出了目标位置。

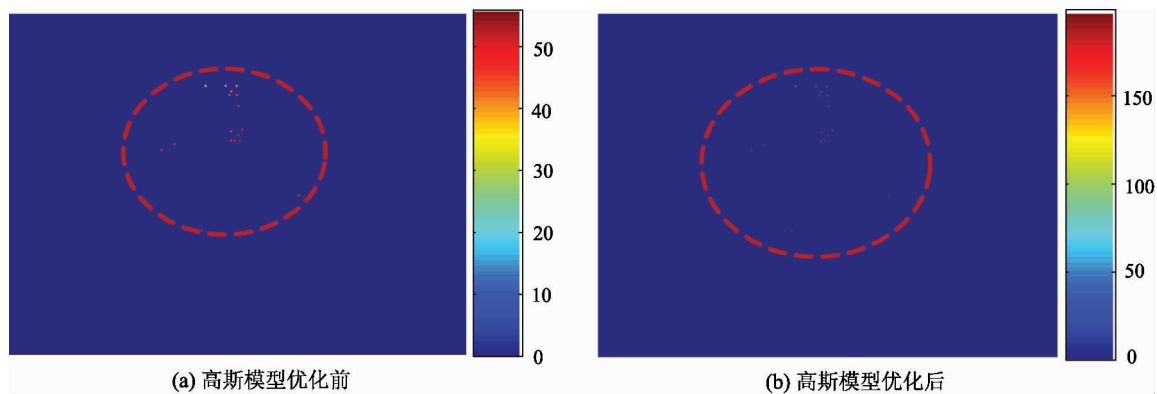


图6 高斯模型优化后前后对比

3 实验及结果分析

本文实验的运行环境为 Intel Core i3-2310 2.1GHz PC, 内存 6G, Win7 64 位操作系统, Matlab 运行平台。

为了验证本文算法的执行速度,利用 Code and Data 数据库的 david_indoor 人脸图像,提取其中第300帧到第399帧的100帧分辨率为 320×240 的图像进行跟踪实验。通过与数据库中的注释位置相比较,在误差允许的范围内,实验算法都能进行有效地

跟踪,结果如图7所示,耗时曲线如图8所示。

其中 MHash 为本文的算法,Hash 是 Yu^[20] 等人利用离散余弦变换实现的感知哈希算法,另外,本实验还与当前先进的一些算法进行跟踪速度比较,其中包括跟踪学习检测(TLD)算法,基于时空上下文学习(spatio-temporal context learning, STC)^[22]的跟踪算法 STC,压缩感知目标跟踪算法 CT。执行速度对比结果如图8所示,可见 MHash 的执行速度相对其他算法速度较快,而且执行的平均帧率较平缓,所以 MHash 算法的整体运算量较小。



图 7 标准库跟踪效果

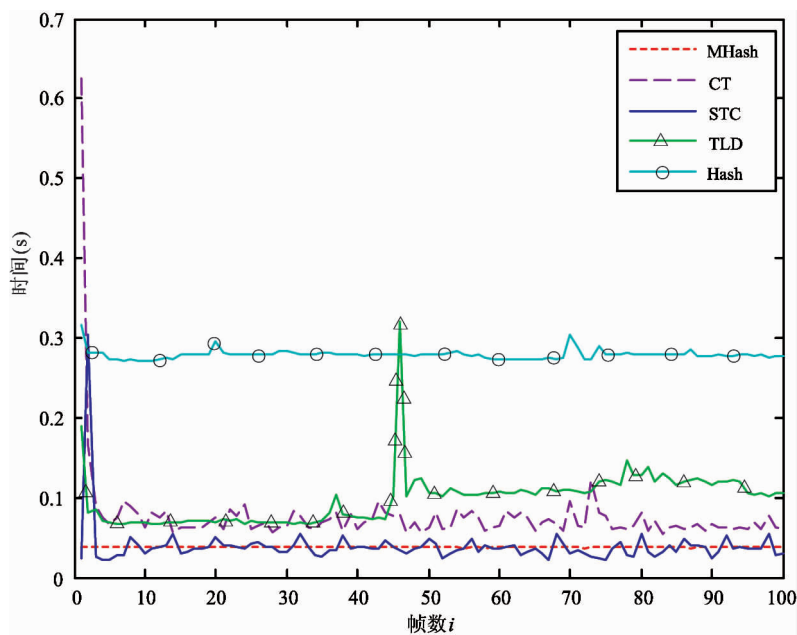


图 8 算法执行耗时曲线

为了验证 MHash 对尺度变化和外形变化的鲁棒性,利用工业摄像头采集运动车辆、人脸的旋转和尺度变化、快速移动和缓慢移动物体环境下图像数据,分辨率为 320×240 。基于 MHash 算法对该 5 组图像数据进行如图 9 所示的跟踪实验,并且与 TLD、STC、CT 和 Hash 目标跟踪算法进行对比。图 9(a)为对处于室外环境的白色和黑色运动车辆跟踪的结

果,由于 MHash 采用的哈希指纹特征对光照变换具有较好的鲁棒性,所以更能适应室外和室内环境的目标跟踪。通过与其他算法相比,MHash 在室外环境中目标跟踪更准确、更稳定。图 9(b)和图 9(c)为对人脸的跟踪结果,图 9(b)中人脸的尺度由小变大,图 9(c)中的人脸进行左右转动,由于采用了正切函数模型对目标窗口进行了平缓的尺度化,可以

看出与其他算法相比较 MHash 对人脸的尺度变化、姿态变化具有较好的鲁棒性。图 9(d)和图 9(e)分别为对快速移动物体和慢速移动物体的跟踪实验结果,图 9(d)中实验者手拿杯子在摄像头范围内缓慢移动,图 9(e)中的动物图案在摄像头范围内快速移动,由于 MHash 采用了双线性内插值提取哈希指纹

特征,有效地降低了运算量,提高了算法的执行速度,增加了目标跟踪的实时性,所以 MHash 对快速移动物体和慢速移动物体均具有较好的鲁棒性。与 Hash 算法相比较 MHash 具有较好的跟踪效果,但与 TLD、STC、CT 算法相比较由于杯子表面纹理特征较少,所以 MHash 对其跟踪效果有待提高。



图 9 算法跟踪效果对比

4 结 论

本文针对光照变化、尺度变化和算法实时性不高的问题提出了一种基于多尺度感知哈希特征的目标跟踪算法。利用正切函数对扫描窗口进行尺度化,采用多尺度窗口对待检测目标进行检测,对目标的尺度变化具有较好的鲁棒性。采用双线性内插值法提取感知哈希特征,有效地降低了算法的计算量,使本文算法实时性更强。同时结合高斯分布和汉明距离对扫描窗口和目标窗口进行匹配评价,有效地

突出了目标位置,滤掉了其他扫描窗口的干扰。实验表明 MHash 算法具有跟踪准确、实时性强,对光照变化、尺度变化鲁棒性强的特点。能够很好地适应室内和室外环境具有较高的实用价值。

参考文献

[1] Chen L, Wei H, Ferryman J. A survey of human motion analysis using depth imagery [J]. *Pattern Recognition Letters*, 2013, 34(15):1995-2006
[2] Lara O D, Labrador M A. A Survey on human activity recognition using wearable sensors[J]. *IEEE Communications Surveys & Tutorials*, 2013, 15(3):1192-1209
[3] Jiang Q. Research of multiple-instance learning for target

- recognition and tracking[J]. *Eurasip Journal on Embedded Systems*, 2016, 2016(1):1-6
- [4] Wang S, Fowlkes C C. Learning optimal parameters for multi-target tracking with contextual interactions[J]. *International Journal of Computer Vision*, 2016, 122:1-18
- [5] Rautaray S S, Agrawal A. Vision based hand gesture recognition for human computer interaction: a survey[J]. *Artificial Intelligence Review*, 2015, 43(1):1-54
- [6] Li K, He F Z, Chen X, et al. Real-time object tracking via compressive feature selection[J]. *Frontiers of Computer Science*, 2016, 10(4):689-701
- [7] 付巍, 郑宾. 基于改进的粒子滤波的静电目标跟踪算法[J]. 高技术通讯, 2014, 24(2):138-143
- [8] 邵万开, 石澄贤. 动态背景下运动目标检测与跟踪研究[J]. 计算机测量与控制, 2016, 24(8):52-55
- [9] 蓝龙, 张翔, 骆志刚. 视觉目标跟踪现状与发展[J]. 国防科技, 2017, 38(05):12-18
- [10] 修春波, 魏世安, 万蓉凤. 二维联合特征模型的自适应均值漂移目标跟踪[J]. 光电子·激光, 2015(2):342-351
- [11] 郇二洋, 李睿. 基于自适应特征融合的粒子滤波目标跟踪算法[J]. 计算机科学, 2015, 42(2):316-318
- [12] Zhang W Z, Ji J G, Jing Z Z, et al. Adaptive Real-Time Compressive Tracking[C]. In: Proceedings of the International Conference on Network and Information Systems for Computers, Wuhan, China, 2016:236-240
- [13] Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2012, 34(7):1409
- [14] 汪鸿翔, 柳培忠, 骆炎民, 等. 单目标跟踪算法综述[J]. 海峡科学, 2016(7):28-32
- [15] 罗海波, 许凌云, 惠斌, 等. 基于深度学习的目标跟踪方法研究现状与展望[J]. 红外与激光工程, 2017, 46(5):6-12
- [16] Dicle C, Camps O I, Szaiaier M. The way they move: tracking multiple targets with similar appearance[C]. In: Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 2014. 2304-2311
- [17] Zhou H, Yuan Y, Shi C. Object tracking using SIFT features and mean shift[J]. *Computer Vision & Image Understanding*, 2009, 113(3):345-352
- [18] Hadmi A, Puech W, Said B A E, et al. A robust and secure perceptual hashing system based on a quantization step analysis[J]. *Signal Processing Image Communication*, 2013, 28(8):929-948
- [19] Pei B, Wang S Z, Li-Jia N I. Content-based image retrieval using perceptual image Hashing[J]. *Journal of Shanghai University*, 2012, 18(4):335-341
- [20] Yu S, Jiang Z. Visual tracking via perceptual image hash from a mobile robot[C]. In: Proceedings of IEEE International Conference on Information and Automation, Lijiang, China, 2015. 1612-1616
- [21] Zhang Y, Zhu C, Bres S, et al. Encoding local binary descriptors by bag-of-features with hamming distance for visual object categorization[C]. In: Proceedings of the European Conference on Advances in Information Retrieval, Moscow, Russia, 2016. 630-641
- [22] 张庆刚, 陈玮. 融合感知哈希特征的时空上下文目标跟踪算法[J]. 信息技术, 2017(4):144-147

The research of a target tracking algorithm based on multi-scale hash feature

Zhang Liguo, Wang Song, Jin Mei, Cui Liyang

(Measurement Technology and Instrumentation Key Laboratory of Hebei Province,
Yanshan University, Qinhuangdao 066004)

Abstract

Target tracking and detection are studied. Considering that the complex application environments bring the problems of low robustness to illumination variation, data loss due to the change of target scale, and low detecting speed during real-time tracking, a new object tracking algorithm using multi-scale perceptual Hash feature, called the MHash algorithm for short, is proposed based on the characteristic that Hash fingerprint features are robust to illumination changes. The algorithm uses multi-scale scanning windows to detect objects and reduces the dimension of scanning window to access perceptual Hash fingerprint characteristics with the bilinear interpolation method. And it combines the Hamming distance with the Gaussian distribution method to evaluate scanning windows. The window of the highest score is as the target window. The method is applied to tracking and detection of face, sports cars and fast and slow moving objects, and the experiments are analysed. The experiments show that this algorithm is robust against illumination changes and target scale change. It can be effectively used for long-term target tracking.

Key words: object tracking, multi-scale, perceptual hash, Hamming distance, feature extraction