

基于区域特征匹配的扩展目标高精度跟踪

张建伟

成都大学电子信息工程学院, 四川 成都 610106

摘要 为了能够更精确的跟踪扩展目标,提出了一种基于最大稳定极值区域(MSER)特征匹配的高精度扩展目标跟踪方法,该方法通过对扩展目标图像中的一个区域进行抽样式多阈值二值化处理提取出MSER,再通过对MSER面积变化率条件的判断保证质心提取精度,并与上一帧MSER进行相似度匹配找到下一帧的MSER,将其质心作为跟踪坐标。实验结果表明该方法对于存在MSER特征的扩展目标具有更精确和更稳定的跟踪能力,在面积变化率一定的条件下,其跟踪精度不低于0.1 pixel,且计算时间也减少了约3/4。

关键词 图像处理;目标跟踪;匹配;质心;亚像素

中图分类号 TP391.41 **文献标识码** A

doi: 10.3788/LOP52.051004

High-Precision Extended Object Tracking Based on Region Feature Matching

Zhang Jianwei

School of Electronic and Information Engineering, Chengdu University, Chengdu, Sichuan 610106, China

Abstract In order to track extended object more precisely, a high-precision extended object tracking method is proposed based on maximally stable extremal regions (MSER) feature, according to binarization with sampling multi-threshold, the MSERs are detected in a region of extended object, if area change rate of MSER meets centroid precision condition, the MSER is picked out. And MSER in next frame are extracted and matched with the MSER in previous frame, the centroid of the MSER matched is next tracking position. Experimental results show that the proposed tracking method can improve the stabilization and precision of extended object tracking, if area change rate of MSER meets the certain condition, the tracking precision is better than 0.1 pixel, and the time consumption reduces by 3/4.

Key words image processing; object tracking; matching; centroid; sub-pixel

OCIS codes 100.3008; 100.2960; 100.4999

1 引言

目标跟踪技术是基于图像处理、模式识别及视频分析等多学科相交叉的一个计算机视觉分支,其主要目的就是在连续的图像序列中能够不断捕获到同一目标的位置。从目标种类来分主要有:微弱目标、小目标、大目标以及扩展目标,针对不同种类的目标有着不同的跟踪方法,而这些方法之间既有区别也有共同之处,例如:对小目标常用的方法有:卡尔曼滤波法、粒子滤波法^[1-2]等等;对于大目标的跟踪方法除了用到滤波预测的方法外,还要利用更丰富的目标特征进行匹配^[3-4];而对于扩展目标其最大的特点是目标范围已经在视场内占据比例较大或者充满甚至超出视场,这种情况就需要对其进行更高精度的跟踪和定位,这在军事上是具有十分重要意义的。

跟踪方法主要针对扩展目标的高精度定位系统,随着科学技术的发展,单独一个摄像机跟踪目标的时代已经逐渐不能满足应用需求了,为了提高跟踪目标的精度,一般采用两个摄像机分别用于粗视场和精视场,在粗视场中,视场较大,主要实现对目标的粗略跟踪,然后再利用精视场进一步提高跟踪定位的精度,这时就需要使用扩展目标的跟踪方法。扩展目标的跟踪方法能够实现更高精度的跟踪,通常可以达到亚像素级

收稿日期: 2014-11-17; 收到修改稿日期: 2014-12-27; 网络出版日期: 2014-05-01

基金项目: 成都大学校基金重点项目(2013XJZ07)

作者简介: 张建伟(1980—),男,博士,讲师,主要从事计算机视觉、目标跟踪等方面的研究。E-mail: zjw75094@163.com

^[5-6],这是相对普通的跟踪方法最大的优势。

目前,扩展目标的跟踪研究相比其他种类目标的跟踪与研究要少一些,但也还是取得了一定的进展。通常情况下,为了达到亚像素级的跟踪精度,扩展目标的跟踪定位过程是采用先特征提取,然后跟踪提取的特征解算亚像素的跟踪位置的方法。在该过程中,特征的选择是核心,常见的特征有点特征^[7]、线特征^[8]和区域特征^[9]。之前一些研究提出了采用直线的交点并利用其进行目标跟踪得到亚像素级的位置。但这种方法稳定性较差,另外还有用尺度不变特征变换(SIFT)跟踪得到亚像素^[10],但是其复杂度又太高。实际上求解亚像素实用又稳定的方法是求质心,质心求解的前提是得到一个稳定且完整的闭合区域,对于细节特征丰富的扩展目标来讲是非常有利的,一般的分割方法很难连续帧中得到完整稳定的同一闭合区域,例如常用的方法有种子填充法、分水岭法^[11-13]、Ostu法^[14]等都在实际跟踪中缺乏实用性,本文提出一种基于最大稳定极值区域(MSER)^[15]特征的方法来实现扩展目标的高精度跟踪。

2 最大稳定极值区域

由于扩展目标在视场中占据比例较大甚至全部覆盖和超出,拥有像素数量多,细节特征比较丰富,增加了区域特征存在的概率,为了能够更准确、更稳定的提取到扩展目标中的区域特征,这里使用MSER进行扩展目标的区域特征提取。

MSER最早由Matas提出^[15],在介绍MSER前,先介绍一下极值区域(ER)的概念,如图1所示。设图像 I 中存在集合 $Q \subset D$ (D 为4连通区域)是区域中的边界像素,对于像素 $p \in Q$ 和 $q \in D$,如果满足 $I(q) < I(p)$ 或 $I(q) > I(p)$,则该区域被称为ER。换言之,就是存在一个区域其边界像素都比其内部区域的所有像素灰度值小或大,那么这个区域就是极值区域。但是极值区域在连续帧中的变化非常不稳定,实用性较差。为了进一步提高极值区域的稳定性,便引入了MSER最大稳定极值区域,与ER极值区域的最大区别就在于其具有更好的稳定性。

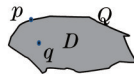


图1 极值区域

Fig.1 Extremal region

按照Matas的方法,要提取MSER主要分以下几步:1)对图像进行二值化,该二值化不是采用唯一阈值,而是从0到255每个灰度级都要作为阈值进行一次二值化;2)提取每个阈值情况下的极值区域,如果在二值化阈值 $[a \ b]$ ($a, b \in [0 \ 255]$)的范围内,某一极值区域面积未发生变化或者变化很小,将该区域认为是备选最大稳定极值区域;3)将设定 $|a-b|$ 的值大于某一数值时就认为该区域为最大稳定极值区域。

最大稳定极值区域主要有以下特性:

- 1) 图像的灰度具有仿射不变性;
- 2) 由于在一定的阈值变化范围内区域边界未发生变化,故其具有较好的稳定性;
- 3) 多尺度的可检测性。

3 基于MSER的扩展目标跟踪

3.1 改进的MSER

MSER并不能直接用于高精度扩展目标跟踪,还存在计算时间和计算精度的问题,为了保证实时性和精确性,这里将抽样二值化和面积变化率的概念引入到MSER中。

首先是抽样二值化,主要考虑计算量上优化,上面提到在 $[0 \ 255]$ 范围内进行多次二值化操作,将其进行简化,对 $[0 \ 255]$ 进行间隔 $s=2$ 或 $s=3$ 的采样,这里称其为抽样二值化,以减少运算量,为了验证抽样二值化的正确性,将抽样二值化与连续二值化的提取结果进行对比。图2为当 $T=1, 3, 7 \dots 255$ 时得到的二值化图像,图3为 $T=1, 2, 3 \dots 254, 255$ 时得到的二值化图像。为了方便比较,这里以图2(a)的箭头所指目标区域作为MSER测试区域,在间隔 $s=2$ 抽样二值化时MSER的阈值范围为 $[107 \ 127]$,而在连续二值化时MSER的阈值范围为 $[107 \ 128]$,因此采样导致了阈值范围的减小,这主要是由采样误差造成的,理论上抽样间隔越大,误差越大,但误差不会超过采样间隔 s 。

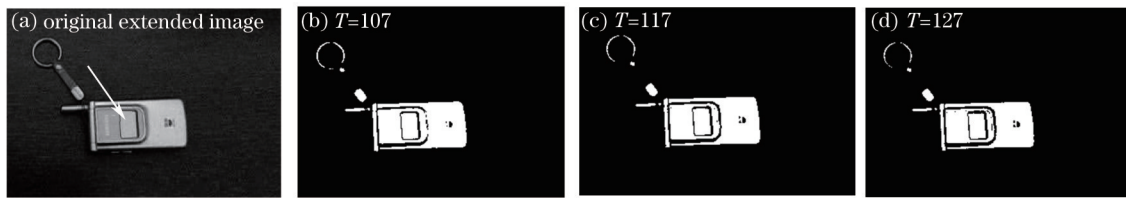


图2 采样二值化得到的极值区域

Fig.2 Extremal region by sampling binarization

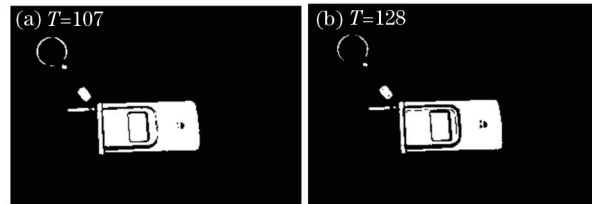


图3 连续二值化采样得到的极值区域

Fig.3 Extremal region by continuous binarization sampling

其次是面积变化率,主要是为了保证计算精度,计算一个区域的质心位置精度需要保证区域面积的像素总数量以及边缘像素变化的数量,这是因为同一区域像素数的变化直接影响到质心的精度量级。把提取同一区域在 $[a \ b]$ 阈值区间内的平均变化像素数与平均总像素数的比称为面积变化率。为了更好的量化面积变化率,以正方形区域为标准对区域面积像素数量的变化对质心提取精度的影响进行分析。此处所采用的测试图如图4中的黑色正方形区域。具体数值如表1所示。



图4 质心精度与面积变化率的关系测试图像

Fig.4 Test image for relationship between centroid precision and area change rate

表1 区域像素总数与质心精度的关系

Table 1 Relationship between centroid precision and number of region pixel

Number of region pixel	2601	1681	961
Change number of region edge pixel	5	5	5
Area change ratio	0.002	0.003	0.005
X-position precision	0.032	0.041	0.057
Y-position precision	0.038	0.043	0.054
Change number of region edge pixel	10	10	10
Area change ratio	0.004	0.006	0.01
X-position precision	0.044	0.073	0.082
Y-position precision	0.049	0.076	0.086

由表1可知,区域面积900以上,边缘像素变化小于10 pixel的情况下,计算的质心与正方形区域的实际质心相比精度在0.08 pixel左右,误差高于0.1 pixel,换句话说在面积变化率 $\delta=0.01$ 的时候, x 和 y 方向精度可以接近0.08 pixel,同样从表中可以看到, δ 接近0.005、0.02的时候, x 和 y 方向精度可以接近0.05 pixel和0.03 pixel,其他的面积变化率对应的精度可以参考同样的方法计算,表1没有再给出。

如果设 e 为MSER的边缘像素平均变化数, n_i 为MSER总像素数, t 为此时选择的二值化阈值, $t \in [a, b], (b > a)$ 为MSER阈值范围。面积变化率 δ 可以表示为

$$\delta = e / (n_a + n_{a+1} + n_b) / (b - a + 1). \quad (1)$$

质心的计算结果才能满足一定精度,由(1)式可知,当边缘像素平均变化数 $e=0$ 时,也就是面积变化率为0的时候,误差为0;当 e 不为0时,也可以按照表1所述方法计算出相应的定位精度。

3.2 扩展目标跟踪方法

利用 MSER 进行跟踪首先被 Donoser 等^[16]提出,但主要是利用 MSER 进行稳定跟踪,而未涉及到精度。以改进的 MSER 为基础,加上对区域面积变化率的精度判断就可以对扩展目标实现精确跟踪。基于 MSER 的扩展目标跟踪步骤如下:

1) 通过操纵杆或者其他锁定设备在视频序列的第一帧中锁定扩展目标的位置。以锁定位置为中心选取一个边长为 l 的正方形区域,通常边长选取标准为

$$l = \min[m, n]/4, \quad (2)$$

式中 m 为视频分辨率的长, n 为视频分辨率的宽,当选取正方形区域超出图像分辨率时,以到达图像边界的实际长度为准。

2) 对于所选中的区域进行抽样二值化处理,其中每一次二值化的图像需要进行区域填充得到相应的极值区域。通过调整二值化的阈值范围 $[a, b]$,使得该极值区域满足当阈值范围 $|a-b| \geq t$ 时(一般 $t=20$),面积变化率符合(1)式,就认为该区域为满足跟踪精度条件的 MSER,在该过程中也可能同时检测到不止一个满足条件的 MSER,那么选择面积变化率最小的 MSER,并以此 MSER 作为模板,计算其质心作为初始跟踪位置并进入下一步,如果不能得到满足条件的区域将返回重新选择区域。

3) 由于相邻两帧变化很小,可以在下一帧中以上一帧的坐标为中心选取一个正方形区域完成 MSER 的检测,并与上一帧 MSER 进行匹配,选择序贯相似性检测算法(SSDA)为

$$d(x, y) = \sum_{x=0}^{p-1} \sum_{y=0}^{p-1} |W_0(x, y) - W_1(x + x_0, y + y_0)|, \quad (3)$$

式中 $W_0(x, y)$ 为基准匹配图像模板, $W_1(x, y)$ 为实时待匹配图像, $p \times p$ 为 MSER 外接正方形区域大小, $d(x, y)$ 为相似度, (x_0, y_0) 为基准图像在实时图像中的偏移值。

由于匹配的模板都是二值化模板而非灰度模板,计算量并不大,相似度 $d(x, y)$ 越小表明两个区域越相似,

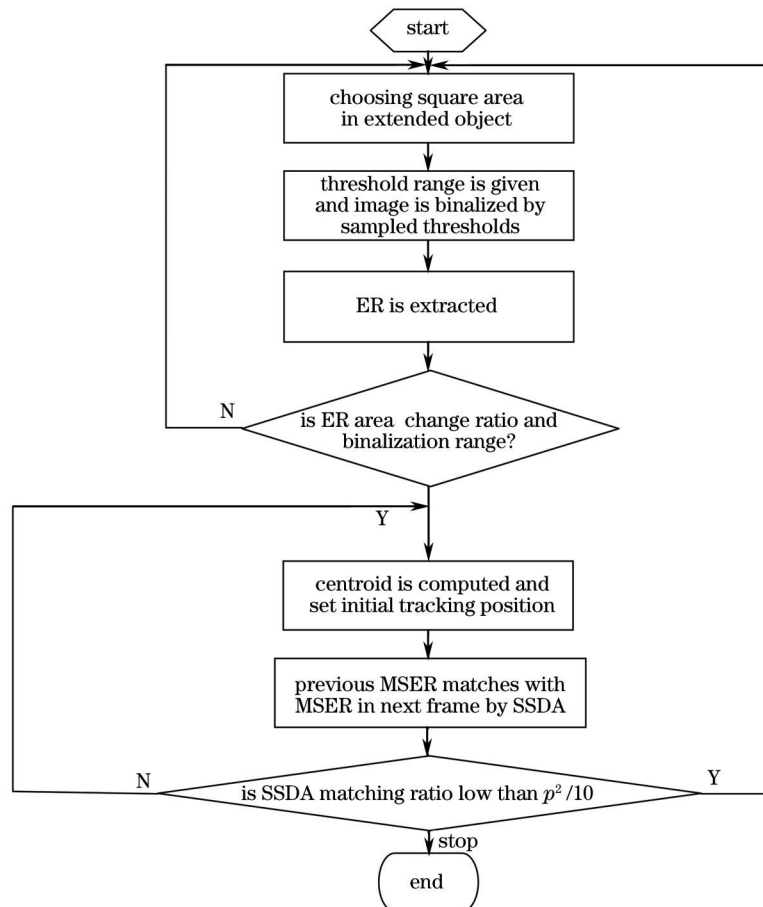


图5 基于 MSER 的扩展目标跟踪流程

Fig.5 Diagram of extended object tracking based on MSER

一般取 $d < p^2/10$, 就认为 MSER 匹配正确, 计算该区域的质心作为该帧的跟踪坐标。其流程图如图 5 所示。

同理, 后续帧可以重复步骤 2) 和 3), 在这个过程中, 若 SSDA 匹配度 $d > p^2/10$ 时进行模板刷新^[17], 使用当前帧的 MSER 区域替换掉第一帧的模板作为最新模板, 然后重复步骤 2) 和 3)。

4 实验结果

利用最大稳定极值区域特征实现对扩展目标的高精度跟踪是一种实现亚像素级定位的过程, 为了验证该基于 MSER 特征匹配算法的跟踪效果, 选取了两段扩展目标视频序列进行了仿真验证。第一段视频中的目标为不同角度拍摄的手机, 由图 6 可知, 白色的方框为初始选择的检测区域, 二值化过程中可以提取到不止一个稳定区域而是四个区域分别为 a, b, c, d , 而在 MSER 的提取时, 从二值化两端的阈值图像来看, b 区域不是一个闭合区域, c 与 b 区域从连通变闭合其稳定程度较差, 而 d 区域太小, 为了保证精度通常会选择区域较大的 a 区域作为匹配模板。 a 区域的阈值范围如图 6 所示, $|a-b|=|127-107|=20$ 满足条件。另外统计得到其最大极值稳定区域像素数平均像素数为 828, 边缘变化平均像素数为 7.6, 因此根据 (1) 式及表 1, $\delta = 0.0092$, 该值小于 0.01 可以保证其质心的精度达到 0.1 pixel。

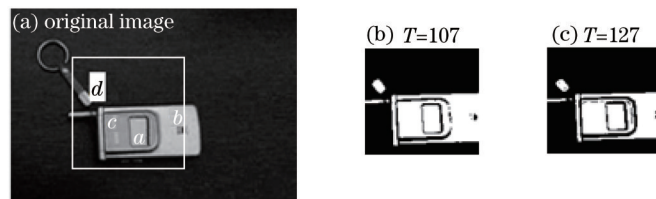


图 6 MSER 选择

Fig.6 Selection of MSER

实际的跟踪效果如图 7 所示。

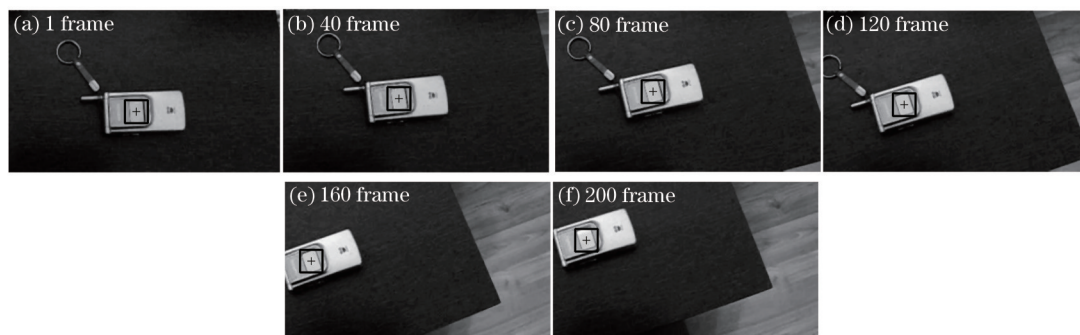


图 7 基于 MSER 的跟踪结果

Fig.7 Tracking performances based on MSER

为了显示基于 MSER 跟踪精度的优势, 图 8 给出它与像素级模板匹配跟踪轨迹的对比, x, y 轴分别为跟踪点 x 和 y 方向在图像中的坐标值。虽然看上去轨迹类似, 但是实际精度差别却很大, 通过图 8(c) 与 (d) 比较可以看出, 模板匹配基本估计都在整数像素的位置, 只能保证精度最小到 1 pixel, 而基于 MSER 跟踪方法, 却可以将跟踪坐标显示到 0.1 pixel。

为了进一步验证该算法的一般性, 又选取了第二段视频, 该视频中的扩展目标为一个模拟正在飞行的飞机, 图 9 黑色方框为所选中区域, 在进行 MSER 提取后, 选取了白色圆圈的黑色区域作为 MSER 的匹配模板, 这主要有两个原因: 1) 该 MSER 区域像素数较多, 平均达到 2535, 边缘变化平均像素数仅为 4.8; 2) 该区域的稳定性较强, 其二值化阈值范围如图 9 所示, $|a-b|=|112-60|=52$ 满足条件。根据 (1) 式 $\delta = 0.0018$, 对照表 1 可以得到该精度在 0.03 左右。

该方法的跟踪精度是伴随着所选择的 MSER 区域变化而变化的, 如图 10 所示, 如果 MSER 区域的面积变化小, 则该跟踪的精度就高, 反之精度就低。在亚像素精度测量与评估方面, 已经有很多方法提出, 但都非常复杂或精度不高^[18-19], 在实验中无需目标标准数据而是以跟踪区域质心为参考对象, 实际上是将正确的质心默认作为了标准数据。

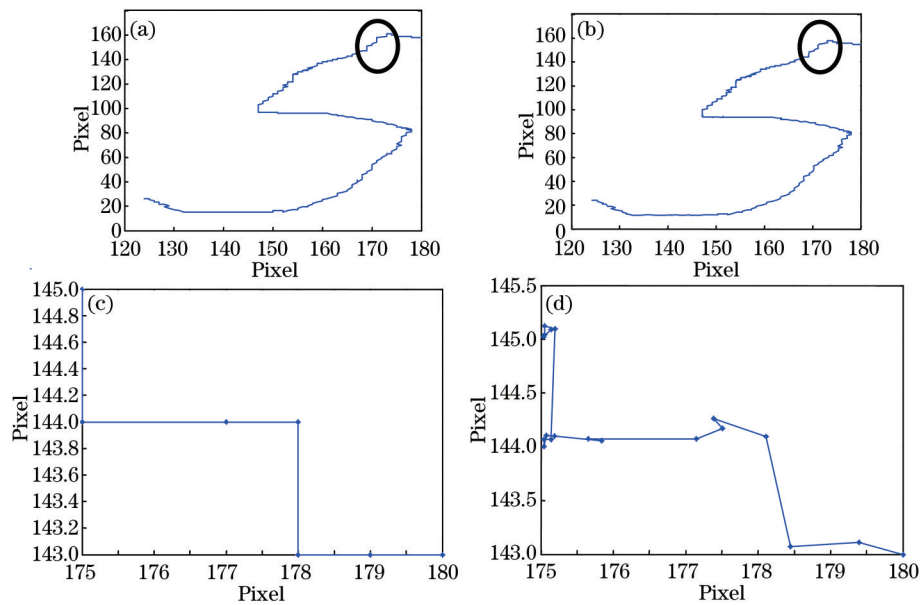


图8 MSER跟踪轨迹与普通模板匹配跟踪轨迹对比。(a)模板跟踪轨迹;(b)MSER跟踪轨迹;
(c)图(a)局部放大;(d)图(b)局部放大

Fig.8 Comparison between MSER tracking and ordinary template tracking. (a) Trace of template tracking; (b) trace of MSER tracking; (c) amplified trace in (a) circle; (d) amplified trace in (b) circle

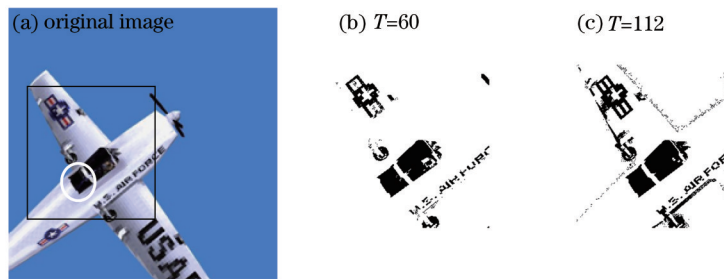


图9 MSER选择

Fig.9 Selection of MSER

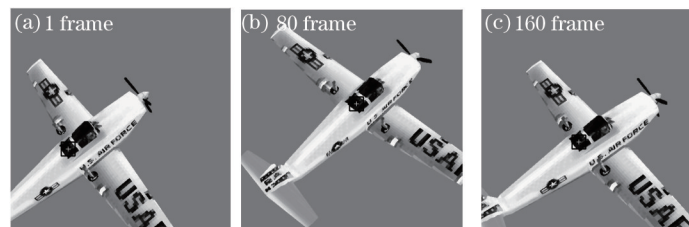


图10 基于MSER的跟踪结果

Fig.10 Tracking performances based on MSER

与其他文献中用的边缘或直线等得到亚像素的方法相比,该方法则具有精度高、方法简单的特点。其他文献中一般都需要特定形状直线^[5-6]或椭圆^[20]进行拟合,而MSER方法不限于目标边缘或者形状使用要灵活得多,在精度上伴随着面积变化率的变化而变化,若面积变化率低于0.01,就可以保证跟踪精度不小于0.1 pixel,当面积变化率达到0.002时,精度可以达到0.03 pixel左右,而文献[5-6]通常精度在0.1 pixel,文献[18]中的精度在0.08~0.03 pixel。

5 结 论

利用MSER特征匹配的方法来进行扩展目标的高精度跟踪,在加入面积变化率的质心精度分析后,实现了对于在精视场中的扩展目标的亚像素级定位精度。实验证明该方法可以很大程度地提高跟踪精度,与像素级模板匹配及其他亚像素定位方法相比,不仅精度高而且复杂度较低,这对闭环云台跟踪系统的精度

提高和实用性有着非常重要的意义;另外在加入抽样二值化的改进后,计算时间也减少了约3/4,在中央处理器(CPU)i7-4500,8GDDR3内存配置的计算机上,仿真软件为Matlab 2009b,每帧的计算时间从1.1 s减少到300 ms,这对于实际嵌入式跟踪系统具有很高的实际价值。并着重研究了基于质心条件的精度判断与分析,但还未对复杂背景、模糊、缺损、遮挡等情况下的算法稳健性进行研究,这也是下一步研究的主要内容。

参考文献

- 1 Wang Xin, Tang Zhenmin. An improved camshift-based particle filter algorithm for real-time target tracking[J]. *Journal of Image and Graphics*, 2010, 15(10): 1507-1514.
王鑫,唐振民.一种改进的基于Camshift的粒子滤波实时目标跟踪算法[J].*中国图象图形学报*, 2010, 15(10): 1507-1514.
- 2 Li Jing, Fan Jiulun. Algorithm for moving object tracking based on Kalman filter[J]. *Application Research of Computers*, 2010, 27(8): 3162-3164.
李晶,范九伦.一种基于卡尔曼滤波的运动物体跟踪算法[J].*计算机应用研究*, 2010, 27(8): 3162-3164.
- 3 Lin Haifeng, Ma Yufeng, Song Tao. Research on object tracking algorithm based on SIFT[J]. *Acta Automatica Sinica*, 2010, 36(8): 1204-1208.
蔺海峰,马宇峰,宋涛.基于SIFT特征目标跟踪算法研究[J].*自动化学报*, 2010, 36(8): 1204-1208.
- 4 Lowe D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- 5 Zhang Jianwei, Zhang Qiheng. An extended object tracking method based on straight-line edge feature matching[J]. *Acta Optica Sinica*, 2009, 29(10): 2826-2831.
张建伟,张启衡.基于直线边缘特征匹配的扩展目标跟踪方法[J].*光学学报*, 2009, 29(10): 2826-2831.
- 6 Liu Qian, Yan Yuzhuang, Huang Xinsheng. Subpixel correlation matching tracking algorithm based on edge feature[J]. *Computer Engineering*, 2011, 37(15): 201-202.
刘倩,闫宇壮,黄新生.基于边缘特征的亚像素相关匹配跟踪算法[J].*计算机工程*, 2011, 37(15): 201-202.
- 7 Gao X, Sattar F, Venkateswarlu R. Multiscale corner detection of gray level images based on log-gabor wavelet transform[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2007, 17(7): 868-875.
- 8 Dong Jing, Yang Xia, Yu Qifeng. Fast line segment detection based on edge connecting[J]. *Acta Optica Sinica*, 2013, 33(3): 0315003.
董静,杨夏,于起峰.基于边缘连接快速直线段检测算法[J].*光学学报*, 2013, 33(3): 0315003.
- 9 Chetverikov D, Matas J. Periodic textures as distinguished regions for wide-baseline stereo correspondence[C]. In *Proc. of the 2nd International Workshop on Texture Analysis and Synthesis*, 2002: 25-30.
- 10 Kupfer B, Netanyahu N S, Shimshoni I. An efficient SIFT-based mode-seeking algorithm for sub-pixel registration of remotely sensed images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 12(2): 379-383.
- 11 Couprie M, Najman L, Bertrand G. Quasi-linear algorithms for the topological watershed[J]. *Journal of Mathematical Imaging and Vision*, 2005, 22(2-3): 231-249.
- 12 Lapeer R J, Tan A C, Aldridge R. Active Watersheds: Combining 3D Watershed Segmentation and Active Contours to Extract Abdominal Organs from MR Images[M]. Berlin: Springer, 2002: 596-603.
- 13 Chen Jiabin, Wang Jigang. Improved watershed segmentation algorithm for medical image[J]. *Application Research of Computers*, 2013, 30(8): 2557-2560.
陈家新,王纪刚.一种改进的医学图像分水岭分割算法[J].*计算机应用研究*, 2013, 30(8): 2557-2560.
- 14 Wang Fangyu, He Xin, Wei Zhonghui, et al.. Automated extraction of projectile in infrared serial images[J]. *Infrared Technology*, 2012, 34(11): 640-643.
王方雨,何昕,魏仲慧,等.红外高炮弹点目标序列图像自动提取[J].*红外技术*, 2012, 34(11): 640-643.
- 15 Matas J, Chum O, Urban M, et al.. Robust wide baseline stereo from maximally stable extremal regions[J]. *Image and Vision Computing*, 2004, 22(10): 761-767.
- 16 Donoser M, Bischof H. Efficient maximally stable extremal region (MSER) Tracking[C]. *Computer Vision and Pattern Recognition*, 2006, 1: 553-560.
- 17 Matthews I, Ishikawa T, Baker S. The template update problem[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(6): 810-815.

- 18 Liu Guodong, Liu Bingguo, Chen Fengdong, *et al.*. Study on the method of accuracy evaluation of sub-pixel location operators[J]. *Acta Optica Sinica*, 2009, 29(12): 3446-3451.
刘国栋, 刘炳国, 陈凤东, 等. 亚像素定位算法精度评价方法的研究[J]. *光学学报*, 2009, 29(12): 3446-3451.
- 19 Yang Yong, Wang Yanlei, Li Ming, *et al.*. Research of high-accuracy digital image correlation measurement system[J]. *Acta Optica Sinica*, 2006, 26(2): 197-201.
杨 勇, 王琰蕾, 李 明, 等. 高精度数字图像相关测量系统及其技术研究[J]. *光学学报*, 2006, 26(2): 197-201.
- 20 Xie Zexiao, Gao Xiang, Cui Jian. Extraction algorithm of circular targets used for mobile three-dimensional measurement[J]. *Chinese J Lasers*, 2013, 40(12): 1208002.
解则晓, 高 翔, 崔 健. 移动式三维测量用圆形标记点提取算法[J]. *中国激光*, 2013, 40(12): 1208002.

栏目编辑: 苏 岑