

子空间模型下的仿射不变目标跟踪

崔雄文^{1,2}, 吴钦章¹, 蒋平¹, 周进¹

(1. 中国科学院光电技术研究所, 四川 成都 610209; 2. 中国科学院大学, 北京 100049)

摘要: 针对目标跟踪过程中目标可能出现的快速变化和严重遮挡等问题, 提出了一种基于新的子空间表示的目标跟踪算法。采用距离不变量对尺度不变特征变换(SIFT)特征点匹配对进行提纯。用提纯后的特征点匹配对, 通过线性拟合得到仿射变化参数。在粒子滤波的理论框架下, 采用快速的迭代算法, 建立目标的主分量(PCA)子空间表示, 结合计算得到的仿射变化参数, 构造有效的目标观测模型完成跟踪。同时, 采用在线学习的方法对 SIFT 特征点和 PCA 子空间进行定时更新。大量实验表明, 提出的算法能快速有效地完成对姿态和形状剧烈变化的目标的精确跟踪。

关键词: 目标跟踪; SIFT 特征点; 线性拟合; 粒子滤波; PCA 子空间; 在线学习

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1007-2276(2015)02-0769-06

Affine-invariant target tracking based on subspace representation

Cui Xiongwen^{1,2}, Wu Qin Zhang¹, Jiang Ping¹, Zhou Jin¹

(1. Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: A new subspace-based tracking algorithm is proposed to deal with rapid changes and severe occlusion of targets during target tracking. Distance invariants are used to select superior pairs of SIFT features. Affine transformation parameters of affine transformation are calculated via linear fitting using these superior pairs. PCA subspace representation of the target is calculated via quick iterations. An effective observation model is constructed to track targets by combining this representation with obtained parameters under the framework of particle filter tracking. And SIFT features and PCA subspace are updated via online learning regularly. At the end, lots of experiments are performed and it is certificated that the proposed algorithm is able to precisely track targets of which gestures and shapes vary rapidly.

Key words: target tracking; SIFT features; linear fitting; particle filter; PCA subspace; online learning

收稿日期: 2014-06-09; 修订日期: 2014-07-13

基金项目: 国家高新技术研究发展计划(G107302)

作者简介: 崔雄文(1988-), 男, 博士生, 主要从事目标探测、识别和跟踪方面的研究。Email: cuixiongwen@126.com

导师简介: 吴钦章(1955-), 男, 研究员, 博士生导师, 主要从事光电测量信息处理及计算机控制管理技术的研究。

Email: qzwu@ioe.ac.cn

0 引言

目标跟踪是计算机视觉的一个研究热点,在红外成像制导、航天测量、宽基线匹配、视频监控等领域有广泛的应用。当前跟踪研究主要集中在基于特征和基于模型的跟踪,以及跟踪和其他视觉问题(检测、识别等)的结合^[1]。

基于特征的跟踪中,局部不变特征基于尺度空间理论构建,具有良好的旋转和尺度不变性^[2-3]。但这类特征存在的不足是,特征提取和特征匹配不稳定性,很难满足精确跟踪的要求。

基于模型的跟踪,根据所使用的目标表观模型,可以分为生成式算法和区分式算法。生成式算法通过在线或者离线的方法建立目标的表观模型,然后搜寻图像中与该模型最为接近的子区域作为目标当前位置。表观模型的建立依赖于目标的描述方法,其中,图像的稀疏表示是有效提高跟踪实时性的一种重要的描述方法。许多文献在建立图像的稀疏表示上做了有效的尝试,如基于压缩感知理论的稀疏表示^[4]和基于子空间的稀疏表示。David A R 等提出了通过矩阵奇异值分解构造目标子空间表示,并采用在线增量学习更新子空间的方法^[5]。Xue,Tang 等采用了结合了平凡模板的目标模板集构造完备的目标子空间,有效解决了部分遮挡的问题^[6-7]。区分式方法将跟踪看作背景和目标在可区分特征下的二类模式分类问题。比较典型的区分式跟踪算法是基于 boosting 的多例学习算法(MIL tracking)^[8]。

上述的跟踪方法都是在 Mean-shift 均值漂移或者粒子滤波的框架下完成。Mean-shift 算法最优寻解采用梯度下降算法,容易陷入局部最优解,因此较多采用粒子滤波的框架^[9]。大多数文献中采样方法为概率随机采样,在目标状态变化较大时,很难实现精确的跟踪。

文中采用粒子滤波框架,结合 SIFT 特征和 PCA 子空间表示,提出了一种能对姿态和形状剧烈变化的目标实现稳健跟踪的新算法。

所提算法流程图如图 1 所示。

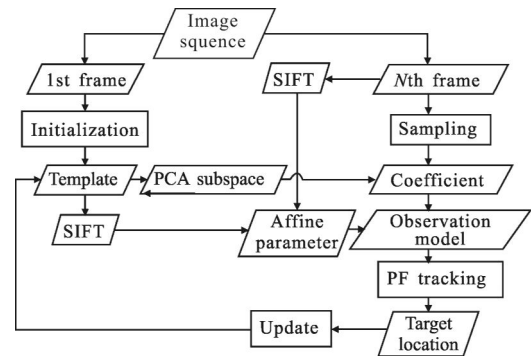


图 1 算法流程图

Fig.1 Flow chart of proposed algorithm

1 SIFT 匹配点对提纯

SIFT 特征的计算步骤如下:

(1) 计算图像 $I(x, y)$ 的尺度空间函数 $L(x, y, \sigma)$:

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

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

(2) 计算尺度空间差分函数 $D(x, y, \sigma)$:

$$D(x, y, \sigma) \approx (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (3)$$

(3) 计算 $D(x, y, \sigma)$ 关于 x, y, σ 的极值点,得到每个特征点对应位置和尺度。

(4) 边缘点剔除和局部非极值约束。

(5) 在由位置和尺度决定的邻近区域内提取梯度信息,进行主方向旋转和归一化操作,得到描述子。利用描述子采用相似度准则即可进行特征点匹配。

当背景复杂且发生变化时,SIFT 特征容易出现误匹配。如图 2 所示,对于 cheese 图像(图 2(a)),虽然视角发生了很大变化,但由于背景没变,因此所有点对均匹配正确;对于 plane 图像(下图),虽然视角变化较小(摄像头都是放在飞机的侧后方),但由于背景变化,导致机头上方出现一个明显的误匹配点对。

刚性目标的运动可以用一个仿射变换来近似。在仿射变换下,线段的长度比短时间内变化较小。因此,对于 n 个匹配点对 $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$,其中, $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)$ 分别为前后帧图像 I_1, I_2 检测到的特征点位置子集,采用以下公式计算每两个匹配点对的距离比值:

$$r(i, j) = r(j, i) = \begin{cases} S & \text{if } \|x_i - x_j\|_2 > T \\ 0 & \text{else} \end{cases} \quad (4)$$

$$S = \|x_i - x_j\|_2 / \min(\|y_i - y_j\|_2, \alpha) \quad (5)$$

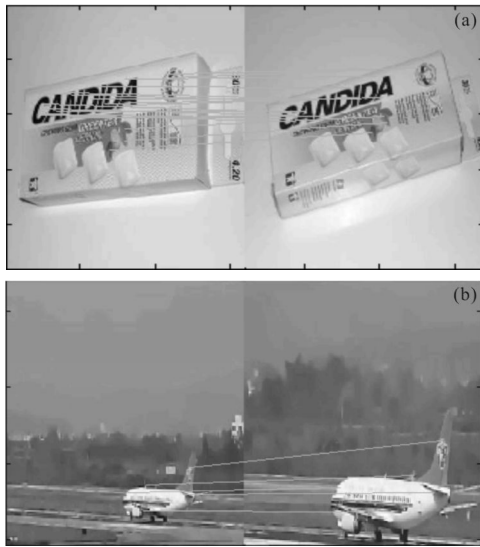


图 2 SIFT 特征点匹配

Fig.2 Matching of SIFT features

$$\delta(i, j) = \delta(j, i) = \begin{cases} 1 & \text{if } \|x_i - x_j\|_2 > T \\ 0 & \text{else} \end{cases} \quad (6)$$

如果特征点太近,前后帧位置提取时存在的微小误差会导致比值的显著错误,因此设置阈值 T 避免这种情况,实验中 T 设定为 5。 α 是为了防止零除的情况,实验中设定为 1。

计算 I_1 中每个点距离比均值为:

$$AR(i) = \sum_j r(i, j) / \sum_j \delta(i, j) \quad i=1, \dots, n \quad (7)$$

式中: $AR(i)$ 表示从匹配点 x_i 连接而成的线段集在前后帧间的长度比均值。对于目标中匹配正确的点对, $AR(i)$ 近似等于目标在前后帧尺度缩放系数;对于匹配错误的点对(从目标点匹配到背景点),这个比值将会出现偏差,予以剔除。

2 线性拟合

$(x_1', y_1'), (x_2', y_2') \dots (x_m', y_m')$ 为提纯后的匹配点对,它们的关系可以表示为:

$$y_i' = Tx_i' + V \quad i=1, \dots, m$$

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \quad V = [S_x \quad S_y] \quad (8)$$

式中: T 为几何变换矩阵; V 为平移矩阵。通过线性拟合的方法求解以下最优化问题得到对应的仿射变换参数 T^*, V^* :

$$(T^*, V^*) = \underset{T, V}{\operatorname{argmin}} \left(\sum_i^m \|y_i' - (Tx_i' + V)\|_2 \right) \quad (9)$$

3 子空间表示下的粒子滤波跟踪

3.1 子空间表示

参考文献[6]中基于 PCA 子空间的增量学习目标跟踪方法,采用矩阵的 SVD 奇异值分解获得目标 PCA 子空间 U 。文中采用同样的方法生成目标子空间。

在遮挡或存在噪声的情况下,仅用非完备子空间 U 无法线性表示目标观测值,因此,文中采用参考文献[7-8]中的平凡模板集来对子空间进行超完备扩充,则目标的一个观测值 Y 在正交基下可以表示为:

$$Y = Uz + e = [U \quad I] \begin{bmatrix} z \\ e \end{bmatrix} \quad (10)$$

式中: e 为平凡模板对应的系数。 Y 的子空间表示等价于求解

$$[Z^* \quad e^*] = \underset{z, e}{\operatorname{argmin}} \left(\frac{1}{2} \|y - Uz - e\|_2^2 + \lambda \|e\|_1 \right) \quad (11)$$

式中: λ 为平凡系数的权值,实验设定为 0.05。参考文献[7-8]采用的 $L1$ 最小化求解方法,计算费时,比较适合小目标的跟踪(如参考文献[7]中 15×20 的小目标和参考文献[9]中的红外小目标)。为提高算法实时性能,参考梯度下降算法,采用误差递减原理迭代求解公式(11):

(1) 初始化 $e_0 = 0$, 迭代次数 $i = 1$;

(2) 采用以下公式计算 Z_i :

$$Z_i = U^T (Y - e_{i-1}) \quad (12)$$

(3) 采用以下公式计算 e_i :

$$e_i = S_\lambda (Y - Uz_{i-1}) \quad (13)$$

$$S_\lambda(x) = \operatorname{sgn}(x) (|x| - \lambda) \quad (14)$$

(4) 若 $\|e_i - e_{i-1}\|_2 < \theta$ 或 $i = K$, 迭代终止;否则 $i = i + 1$, 转到步骤(2)。 K 为设定的迭代次数,在实验中设为 15, θ 为阈值,在实验中设为 0.02。最终得到公式(12)的解 Z^* 和 e^* 。

3.2 粒子滤波跟踪

粒子滤波是序列 Monte-Carlo 采样下基于贝叶斯推理的一种滤波方法^[10],在假设跟踪为 Markov 过程的条件下,给定当前获得的观察值 $Y = \{y_1, y_2, \dots, y_i\}$, 目标的状态量 x_i 可以表示为:

$$p(x_i | Y_i) \propto p(y_i | x_i) \int p(x_i | x_{i-1}) p(x_{i-1} | Y_{i-1}) dx_{i-1} \quad (15)$$

式中: $p(x_i | x_{i-1})$ 为运动方程, $p(y_i | x_i)$ 为观测方程。目标的位置通过对采样粒子进行 MAP 得到。

$$x_i^* = \underset{x_i}{\operatorname{argmax}} p(y_i^i | x_i^i) p(x_i^i | x_{i-1}^i) \quad i=1, \dots, N \quad (16)$$

设目标状态表示为 $x_i = \{x_i, y_i, \theta_i, s_i, \alpha_i, u_i\}$, 分别表示 x, y 方向平移量、旋转角、缩放量、长宽比、倾斜角。运动方程表示为:

$$p(x_i | x_{i-1}) = N(x_i; x_{i-1}, \varphi) \quad (17)$$

式中: φ 为对角方差矩阵。

线性拟合得到的仿射参数表示的目标状态 $x^0(t)$, 近似地估计了目标当前的状态。

对采样点 x_i^i 对应的观测值, 采用参考文献[6]中的方法得到修正观测值 $x_i^{i*} = y_i^i - u$, 其中, u 为 PCA 子空间基向量均值, 采用 3.1 节的方法计算得到对应的 Z_i^* 和 e_i^* 。无遮挡时, 目标真实状态所对应的观测值能用 PCA 子空间表示, 即平凡系数 e_i^* 接近零; 目标出现遮挡时, 真实位置对应的非零平凡系数显著小于错误位置对应的平凡系数^[6]。 y_i^{i*} 在子空间下的重构误差为:

$$\text{Error} = \|y_i^{i*} - UZ_i^*\|_2^2 \quad (18)$$

仅用重构误差来构造观察方程, 并没有考虑公式(11)中平凡系数的加权, 以及正确位置和错误位置两种情况下, 遮挡引起的平凡系数的差异。因此, 文中是在参考文献[7-8]的基础上, 对观察方程进行改进:

$$d(y_i^i | x_i^i) = \exp[-(\|w^i \otimes (y_i^{i*} - UZ_i^*)\|_2^2 + \lambda \sum (1 - w^i))] \quad (19)$$

$$w_j^i = \begin{cases} 1 & \text{if } e_{ij}^* < \delta \\ 0 & \text{else} \end{cases} \quad (20)$$

式中: \otimes 为 Hadamard 积; $\delta = 0.05$ 。

考虑到线性拟合得到的 $x^0(t)$, 观察方程修改为 $p(y_i^i | x_i^i) = d(y_i^i | x_i^i) \exp(-\eta \|x_i^i - x_i^0\|_2)$ (21)

3.3 模型更新

MAP 得到目标的当前状态以及对应的 e^{MAP} 。计算 e^{MAP} 中非零像素所占比例 τ 。若 $\tau < ST$, 则目标没有遮挡或者遮挡极少, 用对应的观测值更新观测值集合; 若 $\tau > LT$, 则目标出现严重遮挡, 不更新观测值; 若 $LT > \tau > ST$, 则用均值 u 取代观测值来更新观测值集合。最后, 定时地采用参考文献[6]中的方法对子空间进行更新, 同时用当前目标区域的 SIFT 点集合替换以前的 SIFT 点集合。

4 实验结果和分析

为验证算法的有效性, 在奔腾双核 2.5 GHz、2 G

内存、Windows XP 的 PC 机上对不同图像序列进行跟踪, 程序平台为: Matlab R2010a, VS2010 和 Opencv2.2.0。

采用最新的压缩跟踪算法(CT)以及多例学习算法(MIL)进行跟踪效果对比。参数设定: 粒子滤波数目为 300, 定时更新间隔为 5 帧。

图 3 和图 4 为姿态剧烈变化的飞机跟踪结果, 点、实线、虚线框分别代表 CT, MIL 和文中算法跟踪结果(下同)。帧数分别为 182 和 150。可以看出: 在目标大小、角度发生剧烈变化时, 前两种算法无法自适应调整跟踪窗口以完成准确跟踪, 跟踪出现偏差; 文中算法采用不变特征点计算变形参数, 提高了跟踪性能。

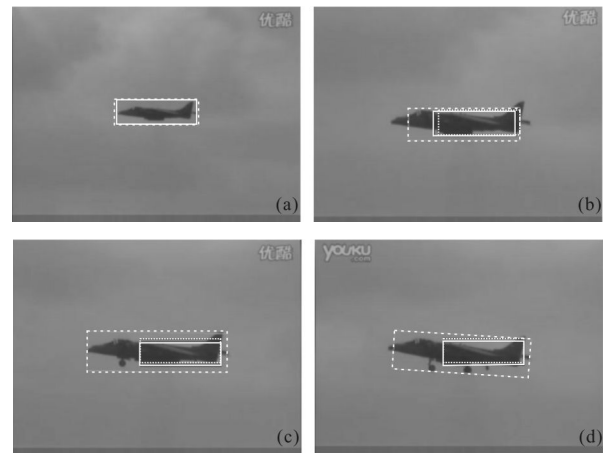


图 3 飞机序列 1 部分跟踪结果(第 1, 49, 128, 170 帧)

Fig.3 Part of tracking results of sequence "plane1"
(in Frame 1,49,128,170)

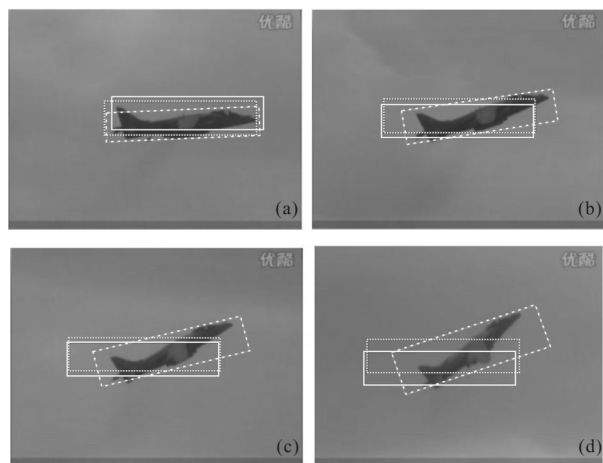


图 4 飞机序列 2 部分跟踪结果(第 46, 77, 100, 143 帧)

Fig.4 Part of tracking results of sequence "plane2"
(in Frame 45,77,100,143)

图 5 为红外小汽车跟踪结果 (序列长度 70 帧), 文中采用改进的 PCA 子空间来描述目标, 相对于 CT 基于压缩感知稀疏描述和 MIL 的 Haar 特征描述, 鲁棒性更强, 跟踪效果较好。

图 6 和图 7 采用标准测试视频“Caviar1”序列和“Girl”序列进行跟踪测试。在“Caviar1”序列中, 目标逐渐缩小且有短暂遮挡, 由于采用了抗遮挡的观测方程和更新方法, 文中算法能稳定准确跟踪目标, MIL 跟踪完全丢失, CT 也出现明显跟踪误差;“Girl”序列中, 目标在垂直于成像平面方向上的运动导致目标区域发生变化, 这种情况下, MIL 和 CT 无法进行完成跟踪。而文中采用了加权更新的方法, 因此, 虽然出现了短暂的跟踪错误(如第 23 帧到第 40 帧), 但能重新准确捕获目标(第 62 帧)。采用视频作者提供的真值和跟踪位置间的欧式距离作为跟踪误差, 两个序列的跟踪结果见图 8。

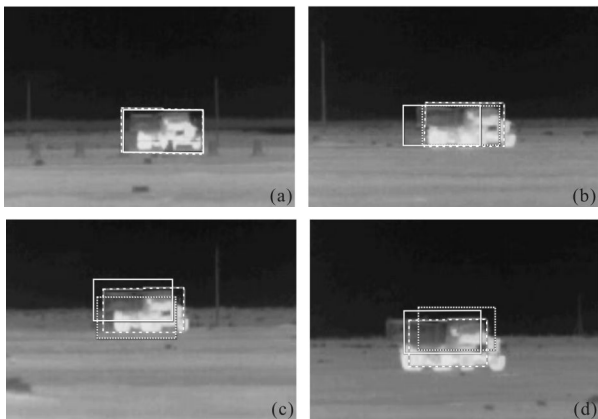


图 5 红外小汽车部分跟踪结果 (第 1,20,40,62 帧)
Fig5 Part of tracking results of infrared sequence "car"
(in Frame 1,20,40,62)

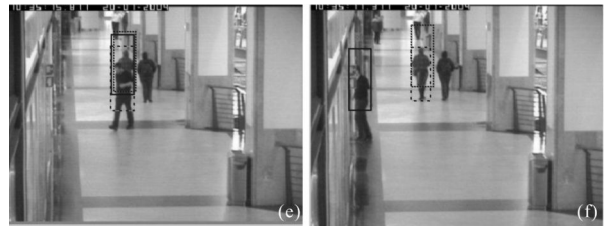


图 6 Caviar1 序列部分跟踪结果 (第 1,50,100,114,136,150 帧)
Fig.6 Part of tracking results of sequence "Caviar1"
(in Frame 1,50,100,114,136,150)

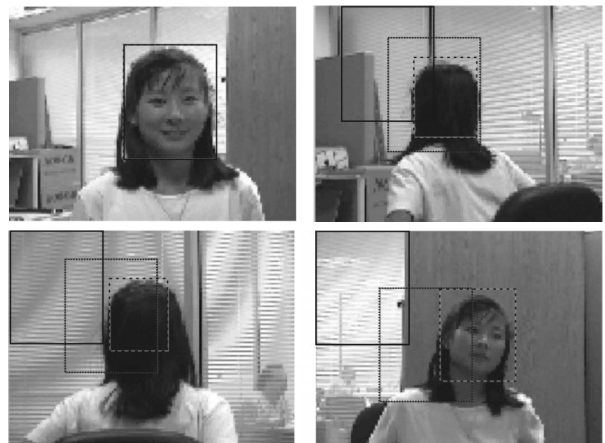


图 7 Girl 序列部分跟踪结果 (第 1,23,40,62 帧)
Fig.7 Part of tracking results of sequence "Girl"
(in Frame 1,23,40,62)

图 8 为 Caviar1 和 Girl 序列跟踪误差。

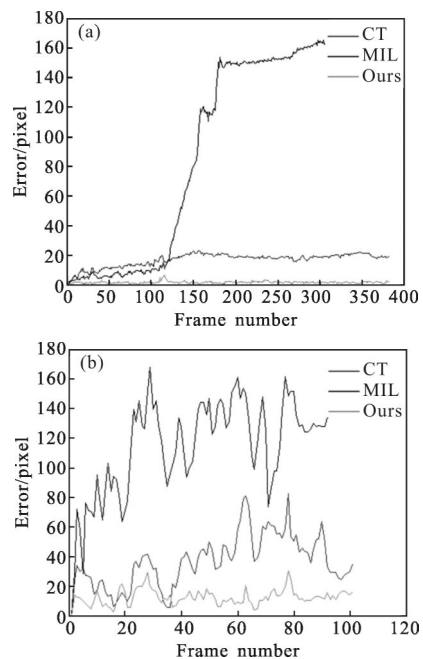


图 8 Caviar1 和 Girl 序列跟踪误差
Fig.8 Tracking errors of sequence "Caviar1" and "Girl"

三种算法的跟踪帧频如表 1 如所,CT 算法采用压缩感知理论构造目标的极度稀疏表达,以及简单的 Haar 特征,因此,实时性最高。参数的选择决定算法的速度,采用文中设定的参数集,在保证稳定可靠跟踪的同时,算法平均跟踪速度为 4.4 帧/s,实时性较高。

表 1 三种算法的跟踪帧频

Tab.1 Frame rates of three tracking algorithms

fps	pl1	pl2	car	Cav	girl	avg
MIL	1.2	1.2	1.2	1.1	1.3	1.20
CT	15.3	16.3	14.2	15.3	15.2	15.26
Ours	5.9	6.5	6.0	1.9	1.7	4.40

5 结论

文中将 SIFT 不变特征和目标的 PCA 子空间描述结合在一起,提出了新的跟踪算法。在目标剧烈运动或者发生遮挡时,能稳定准确地跟踪目标。如何脱离 SVD 算法以建立更有效的子空间表示,以及与检测算法结合以解决视角变化下目标成像变化带来的跟踪误差,是进一步研究所要解决的问题。

参考文献:

- [1] Yilmaz A, Javed O, Shah M. Object tracking:a survey [J]. *ACM Computing Survey*, 2006, 38(4): 1-45.
- [2] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [3] Kong Jun, Tang Xinyi, Jiang Min, et al. IR target tracking based on scale space feature points matching [J]. *Infrared and Laser Engineering*, 2011, 40(11): 2104-2109. (in Chinese)
孔军, 汤心益, 蒋敏, 等. 采用尺度空间特征点匹配的红外目标跟踪[J]. *红外与激光工程*, 2011, 40(11): 2104-2109.
- [4] Kaihua Zhang, Lei Zhang, Ming-husan Yang. Real-time compressive tracking[J]. *Proceedings of European Conference on Computer Vision*, 2012, 3: 864-877.
- [5] David A R, Jongwoo L, Rwei-Sung L, et al. Incremental learning for robust visual tracking [J]. *International Journal of Computer Vision*, 2008, 77(1-3): 1-27.
- [6] Xue Mei, Ling Haibin. Robust visual tracking using L1 minimization [C]//Proceeding of the IEEE International Conference on Computer Vision, 2011: 1257-1264.
- [7] Tang Zhengyuan, Zhao Jiajia, Yang Jie, et al. Infrared target tracking algorithm based on sparse representation model [J]. *Infrared and Laser Engineering*, 2012, 41(5): 1389-1395. (in Chinese)
唐峥远, 赵佳佳, 杨杰, 等. 基于稀疏表示模型的红外目标跟踪算法[J]. *红外与激光工程*, 2012, 41(5): 1389-1395.
- [8] Babenko B, Yang M H, Belongie S. Visual tracking with online multiple instance learning [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2009:983-990.
- [9] Peng Chen, Chen Qian, Qian Weixian, et al. Infrared moving targets tracking under complex ground scene [J]. *Infrared and Laser Engineering*, 2013, 42(6): 1410-1414. (in Chinese)
彭晨, 陈钱, 钱惟贤, 等. 复杂地面场景下的红外运动目标跟踪[J]. *红外与激光工程*, 2013, 42(6): 1410-1414.
- [10] Mikolajczyk K, Schmid C. Detection of local features invariant to affine transformations [C]// European Conference on Computer Vision, 2002: 128-142.