

基于 TLD 的舰船目标跟踪方法研究

齐楠楠¹, 揭斐然¹, 谢 熙², 吴 巍²

(1.中航工业洛阳光电设备研究所 光电控制技术重点实验室, 河南 洛阳 471009;

2.武汉理工大学 信息工程学院, 湖北 武汉 430070)

摘要: 复杂背景下进行舰船目标的跟踪时, 在某些帧可能会有目标丢失。为了克服这个问题, 采用联合检测-学习-跟踪的 TLD 算法。其过程是通过训练一种在线可更新的随机蕨分类器对目标跟踪结果进行检测, 并使用一种基于时空约束的 PN 学习策略对分类器进行学习和更新, 最后融合跟踪得到的结果对目标进行判别和确定。试验结果表明, 该跟踪算法可适用于目标外形改变和遮挡的情况, 鲁棒性强, 识别率高, 误检率低, 同时实时性也较好, 可以满足一般的在线跟踪系统的要求。

关键词: 舰船跟踪; 随机蕨分类器; TLD 算法; 在线学习

中图分类号: TP274; TN713

文献标识码: A

文章编号: 1001-8891(2013)-12-0780-08

Ship Target Tracking Based on Tracking-Learning-Detecting Tactics

QI Nan-nan¹, JIE Fei-ran¹, XIE Xi², WU Wei²

(1.Luoyang Institute of Electro-optical Equipment, Key Laboratory of Optical Electrics Control Technology, Luoyang 471009, China;

2. Wuhan University of Technology, School of Information Engineering, Wuhan 430070, China)

Abstract: When warship targets are tracked in complex background, the targets loss may occur in some frames. In order to overcome the problem, a tracking-learning-detecting (TLD) algorithm is introduced. With the random ferns classifier which is trained online, the detection is performed based on the classification results. PN learning constrained by spatial and temporal features is used to update the classifier. The detection results and tracking results are fused to locate the target in each frame. Finally, experimental result shows that the TLD tracking algorithm has a high recognition rate and a low false detection rate. Benefitting from continuous learning with various target changes in each frame, the TLD algorithm is robust to target appearance changes and occlusion, and has a good real-time performance. The proposed algorithm can meet the requirements of general online tracking system.

Key words: ship tracking, random ferns classifier, TLD algorithm, online learning

0 引言

针对目前跟踪环境多变以及目标跟踪过程中不断变化的情况, 很多学者使用计算机新技术或者其他交叉学科技术, 如小波变换、人工智能, 模糊数学、遗传算法和目标跟踪算法融合。Huan W.^[1], Qin Kun^[2]等人就将生物遗传算法与 mean shift 算法结合, 在复杂背景环境中的跟踪就能取得较好的效果。Nak^[3]等人将 Kalman 滤波预测^[4]和粒子滤波^[5]跟踪算法相结合, 应用到水下目标跟踪项目上, 取得较好的跟踪效果。

一般而言, 长时间的视频目标跟踪需要克服很

多问题, 如相机的位置变化导致视频中前后两帧之间没有绝对紧密的时空联系、候选目标突然消失等。此时, 国外学者 Zdenek Kalal^[6]等人使用跟踪-学习-检测 (Tracking-Learning-Detection: TLD) 联合方案^[7]进行了在线跟踪研究, 取得了很大的进展。本文将该方法引入到舰船目标的跟踪算法中, 以解决长时间舰船目标跟踪过程中的目标尺度变化、局部遮挡所带来的问题。

1 基于 TLD 算法的舰船跟踪框架设计

TLD 舰船跟踪算法框架如图 1 所示, 主要由检测模块、跟踪模块、学习模块构成。跟踪模块对待

收稿日期: 2013-09-06; 修订日期: 2013-10-30.

作者简介: 齐楠楠 (1982-), 女, 硕士, 工程师, 主要研究方向为目标检测和图像跟踪。E-mail: eoei@vip.sina.com。

基金项目: 国家自然科学基金资助项目, 编号: 61273241; 航空科学基金, 编号: 20105179002。

选目标进行跟踪，并同时进行学习和检测；检测模块对跟踪的目标进行判决（只有满足相应分类器的判决才认为是真实目标），并且是在线的、可更新的，可满足实时跟踪的需要；学习模块则通过某些约束条件将通过检测器结果进行变换，并进行检测器的更新，以增加整个系统的鲁棒性。

2 TLD 算法在舰船跟踪的实现

2.1 在线检测器设计

2.1.1 随机蕨分类器构建

在系统中使用随机蕨（random ferns）相关分类器对目标进行检测^[8-10]。随机蕨（random ferns）分类器是由随机森林分类器发展而来^[11]，用来对具有多个特征值的不同目标种类进行数字概率分类，最终使用类似多人投票的方式决定被检测目标的属类。

目标的特征采用了类似 LBP（Local Binary Pattern）特征以及梯度特征，即使用一类 2 bit BP 特征，是检测某一范围图像的梯度信息^[6]，描述如图 2 所示。

目标的每一个特征区域编码结果只占 2 个 bit 位，用来表示该区域上下左右部分之间的颜色深浅关系，如图 2 的特征编码区域中，可先将该区域分为上下两个部分 a_1 和 a_2 ，像素 $I(a_1) < I(a_2)$ 则表示上半部分的颜色较深，2 bit BP 编码中的高位值为 0，反之则为 1；再将区域分为左右两个部分 a_3 和 a_4 ，同样，当 $I(a_3) < I(a_4)$ 时 2 bit BP 编码中的低位编为 0，反之编码为 1。区域上下左右颜色关系确定，则该区域的 2 bit BP 编码也就唯一确定。对于一个检测区域来说，会有很多特征编码区域（ N 个），对应目标区域产生 N 个 2 bit BP 编码。

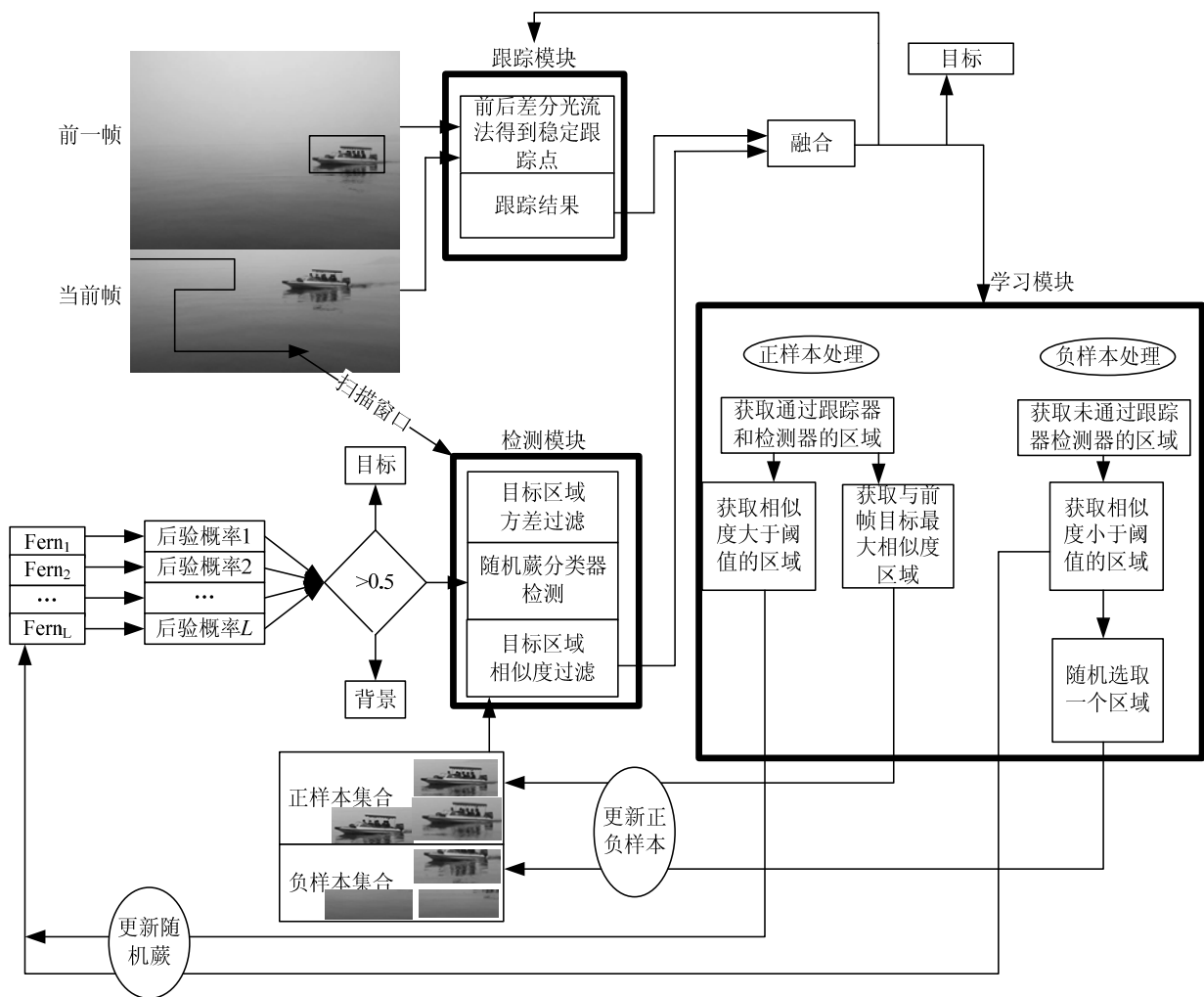


图 1 TLD 算法舰船跟踪流程示意图 Fig.1 Schematic diagram of warship tracking based on TLD algorithm

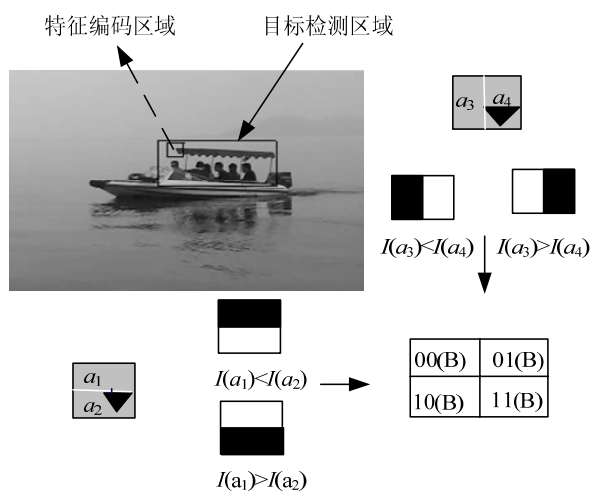


图 2 2 bit BP 特征提取示意图

Fig.2 Diagram of 2 bit BP feature extracting

应用 2 bit BP 编码作为图像的特征对分类器进

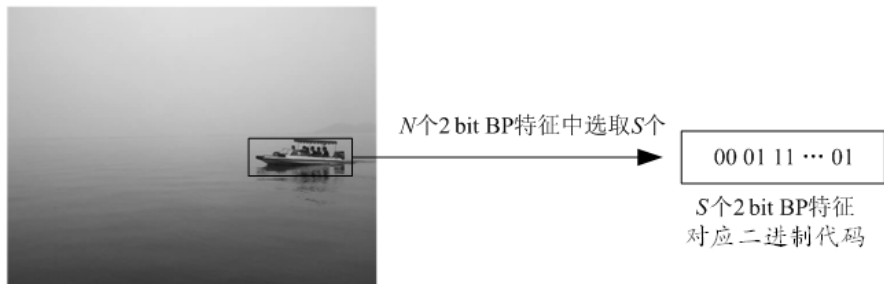


图 3 图像特征随机蕨编码过程图

Fig.3 Encode process of random ferns based on image features

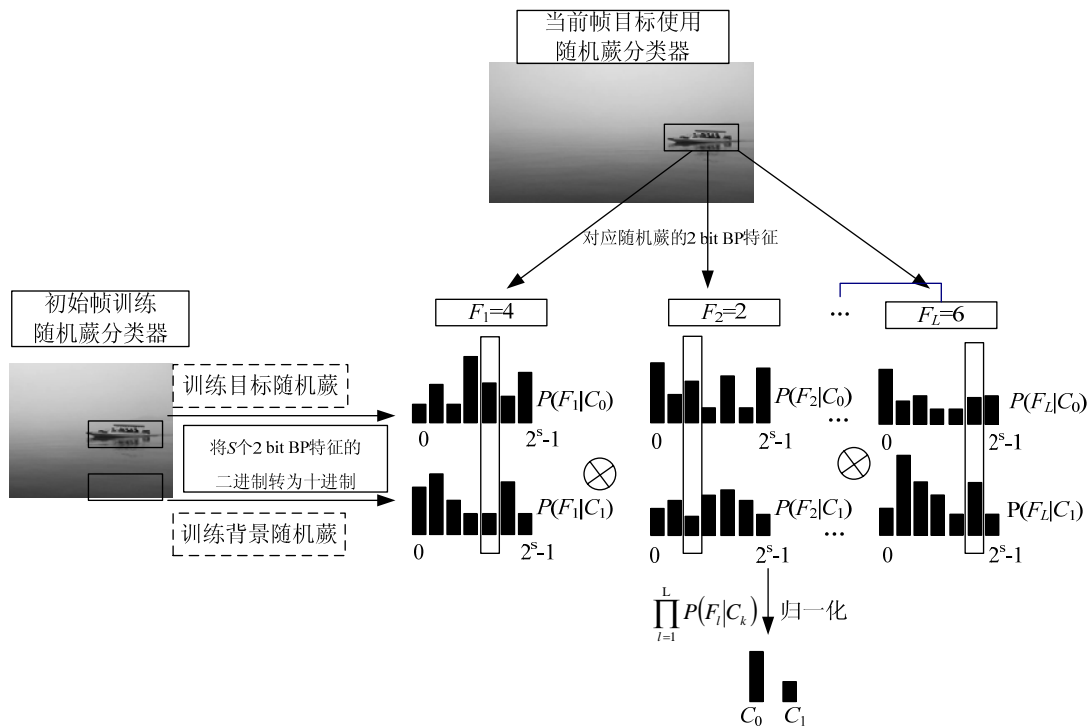


图 4 随机蕨分类器训练及分类示意图

Fig.4 Diagram of training and classifying using random ferns classifier

行训练，从 N 个特征中构建 L 组“特征蕨”，假设每一组含有 S 个随机的 2 bit BP 特征，可以通过目标的初始位置计算随机蕨的先验概率 $P(F_l|C_k)$ ，其中， $C_k(k=0,1)$ 分别为图像中目标区 C_0 和图像中背景区 C_1 ， $F_l(l=1,\cdots,L)$ 表示有第 l 个 2 bit BP 编码特征向量 F 。对未知目标可基于随机蕨的先验概率进行判断，过程如图 3、图 4 所示。

2.1.2 多特征的线检测器设计

整个检测模块的流程如图 5 所示。

在检测过程中，对于被分类的目标区域，依次通过区域方差检测、随机蕨分类器检测以及目标相似度检测，只有都通过这 3 项检测的区域，才认为是真正的目标，一旦不能通过其中一个检测流程，都认为选定区域不是真实目标，同时不再进行其他相关的检测。

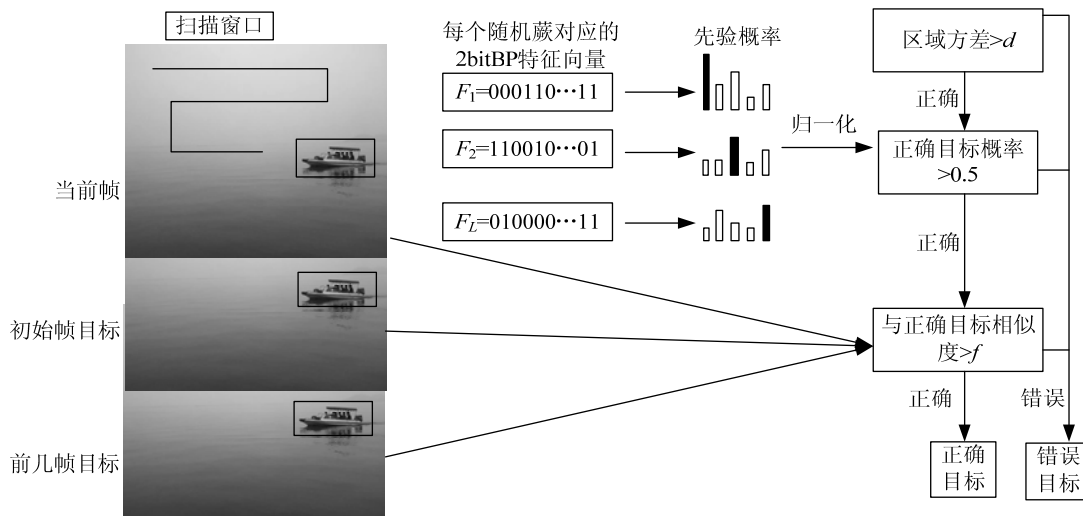


图 5 目标检测流程示意图

Fig.5 Diagram of process for targets detecting

区域方差检测主要是考虑目标图像和背景图像的方差具有差别，在可见光图像中，目标是呈现深色，而周围的背景相对较亮，方差较大，而纯背景则方差很小；在红外图像中，目标的辐射通常与背景不同，表现在图像中也会是方差不同于纯背景图像。

随机藏分类器检测利用目标区域的 2 bit BP 特征组成随机藏分类器，对目标进行分类，当后验归一化概率大于 0.5，认为进入随机藏分类器的区域含有目标。

目标相似度检测使用归一化相关系数（NCC），检测当前目标区域和目标模型（初始帧目标和已经通过检测的目标集合）的相似度，计算如公式 1 所示：

$$S(p_i, p_j) = 0.5(NCC(p_i, p_j) + 1) \quad (1)$$

2.2 跟踪过程中的在线学习

在线跟踪系统中，分类器初始样本很小，得到的分类器不是强分类器，所以必须随着跟踪过程不断地进行学习，优化分类器的检测机制。虽然可以通过目标在帧与帧之间的连续性不断获得新的正负样本，但整个跟踪过程中，会出现分类器结果和正负结构性约束矛盾的情况，此时需要通过其他的约束条件对分类器得到的结果进行修正，这里我们可以使用 P 约束和 N 约束 2 种不同的约束条件对其进行约束和修正^[12]。

PN 约束是可以对分类器进行不断修正，慢慢减小其误差，因此可以在扩大目标的正负样本同时对分类器进行不断学习更新，而不必担心分类器发生错误，保证整个跟踪的正确性。整个学习迭代过程如下：

1) 使用带标签的样本训练分类器，并据此调整相应的正负结构模式。

2) 使用分类器对无标签数据进行分类，找出分类器得到结果与正负结构模式相矛盾的样本。

3) 重新修正矛盾样本的标签并加入到训练集中重新训练分类器。

以上 3 个过程会在整个跟踪过程不断迭代，而且它们的运行可以自动化，不需要人为的控制，这个过程称之为 PN 学习（PN-Learning）^[12]。

受目标的时空约束，认为用于训练分类器的样本可以通过下面两种方式产生：

① 远离目标的区域是背景样本，靠近目标的区域是目标样本。

② 考虑目标可能出现的尺度、旋转、噪声影响，对目标区域进行尺度变化、旋转变换、加噪处理的区域也是正确的目标样本。

2.3 基于光流的在线跟踪

跟踪模块使用前后差分光流法找寻相邻帧之间的稳定跟踪点，对其进行跟踪，相应的跟踪结果会与检测模块中的结果进行融合（其中跟踪器得到的结果权值大于检测器结果的权值），得到被跟踪目标的最终位置，并将融合结果输入到学习模块中进行 PN 学习。

3 实验分析与结论

本次试验均在 CPU 频率为 2.53 GHz，内存为 2 G 配置的 32 位电脑上运行，在 windows7 系统下使用 c 语言编程实现。

跟踪图像使用拍摄的舰船航行视频帧，目标初始尺度为 170×60 像素（尺度变化时初始尺度为 120×120 像素）。所有对比试验中，基于梯度方向

直方图的 mean shift 跟踪算法参数如下：使用 16 个方向的梯度对目标建模，kalman 滤波中观测噪声 $\delta_o=5$ ，预测噪声 $\delta_w=5$ ，尺度更新系数 $r=0.1$ ，位置初始残差 $\mu=4$ 。TLD 跟踪算法参数设置如下：整个图像分为 12×8 个子区域，随机蕨分类器数目 $L=13$ ，每个随机蕨特征数目 $S=6$ 。随机蕨分类器初始阈值为 0.5，相似度初始阈值为 0.55。扫描间隔为 0.05 倍图像长宽，尺度缩放系数为 0.05。

1) 目标尺度变化时跟踪仿真实验

本次试验主要验证目标在发生尺度变化的情况下，使用 TLD 算法能自适应尺度变化。图 6 是此类情况使用 mean shift 跟踪算法得到的效果图，图 7 是使用 TLD 跟踪算法得到的跟踪效果。

如图 6 及图 7 所示，当目标尺度发生变化时，基于 mean shift 算法跟踪和基于 TLD 算法的跟踪都能不断适应目标在运动过程中发生的尺度变化。图 8 是两类算法在目标尺度改变上的对比图（跟踪效果图中可见两类算法都能准确跟踪到目标位置，故不再对跟踪位置进行对比），可以清晰地看出 TLD

算法和 mean shift 算法在尺度收敛上具有相近效果。
2) 目标在帧间存在大位移条件下的跟踪仿真实验

此试验验证目标与前一帧目标偏离巨大时 TLD 算法可以准确跟踪到目标。为了目标巨大位移情况下跟踪情况，将第 23、24、38、39 帧图像里的目标进行较大的位置平移。图 9 是使用 mean shift 跟踪算法得到的效果图，图 10 是使用 TLD 跟踪算法得到的跟踪效果。

在图 9、图 10 中，目标在第 23 帧和第 38 帧发生较大的位置移动，基于 mean shift 跟踪算法在此时通过预测进行目标跟踪，出现了误跟踪（如图 9(b)）；TLD 算法则通过前数帧建立分类器，即使在跟踪模块失败时，也能通过检测模块对全局相似目标进行分类，最终成功跟踪目标。图 11 是巨大位移时 2 种算法的跟踪位置曲线图，图 11(a)可以明显地看出目标在 x 方向发生了剧烈的位置移动，但 TLD 算法仍然能准确地跟踪到目标位置；同样，从图 11(b)中也可以看出，该算法在 y 方向上也适用。



(a)第2帧 (a) No.2 frame (b)第30帧 (b)No.30 frame (c)第60帧 (c)No. 60 frame
图 6 图像序列中目标尺度变化时 mean shift 算法跟踪效果

Fig.6 The tracking results based on mean shift algorithm when the targets' scales in image sequence are changing



(a)第2帧 (a)No.2 frame (b)第30帧 (b) No.30 frame (c)第60帧 (c) No.60 frame
图 7 图像序列中目标尺度变化时 TLD 跟踪效果图

Fig.7 The tracking results based on TLD algorithm when the targets' scales in image sequence are changing

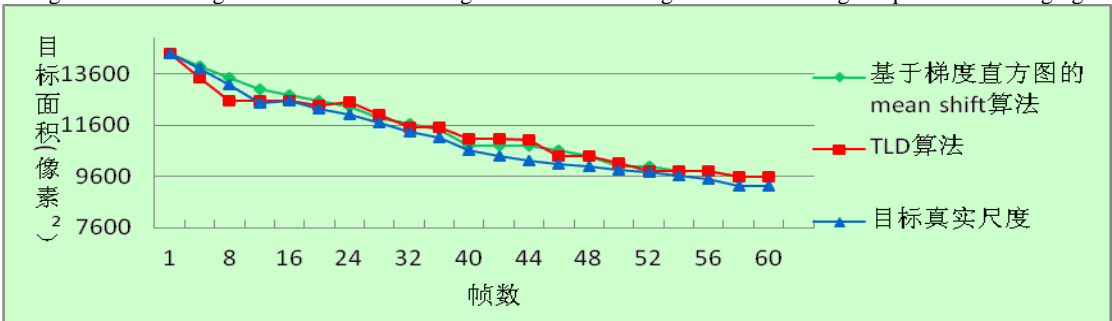


图 8 图像序列中目标尺度变化时性能比较

Fig.8 Diagram of tracking performances when the targets' scales in image sequence are changing

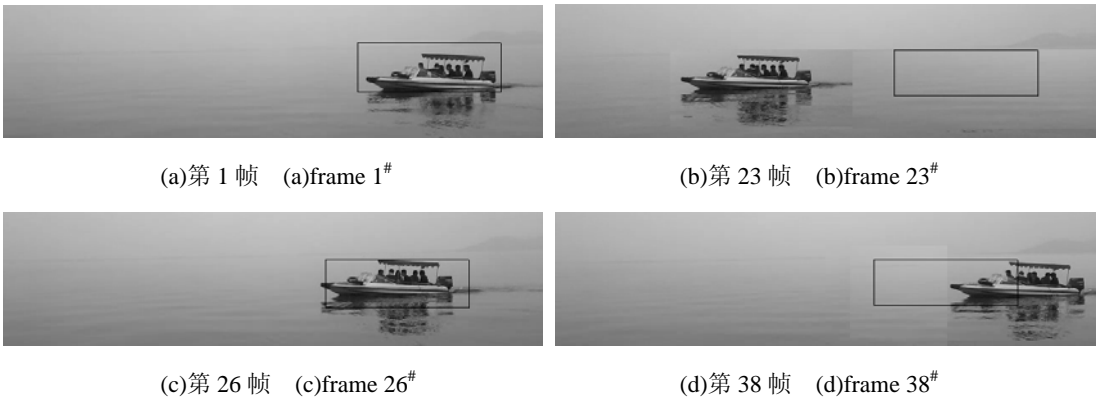


图 9 目标在帧间存在大位移时 mean shift 跟踪效果

Fig.9 The tracking results based on mean shift algorithm when the targets have large displacement in image sequences

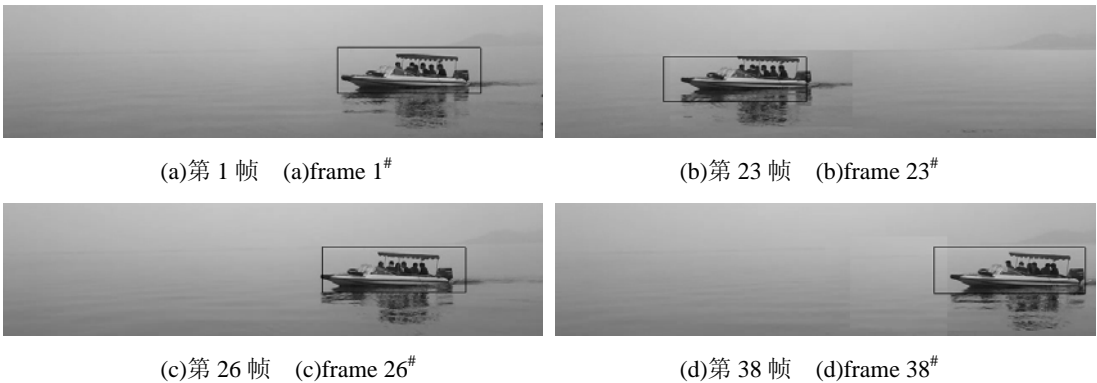


图 10 目标在帧间存在大位移时 TLD 跟踪效果

Fig.10 The tracking results based on TLD algorithm when the targets have large displacement in image sequences

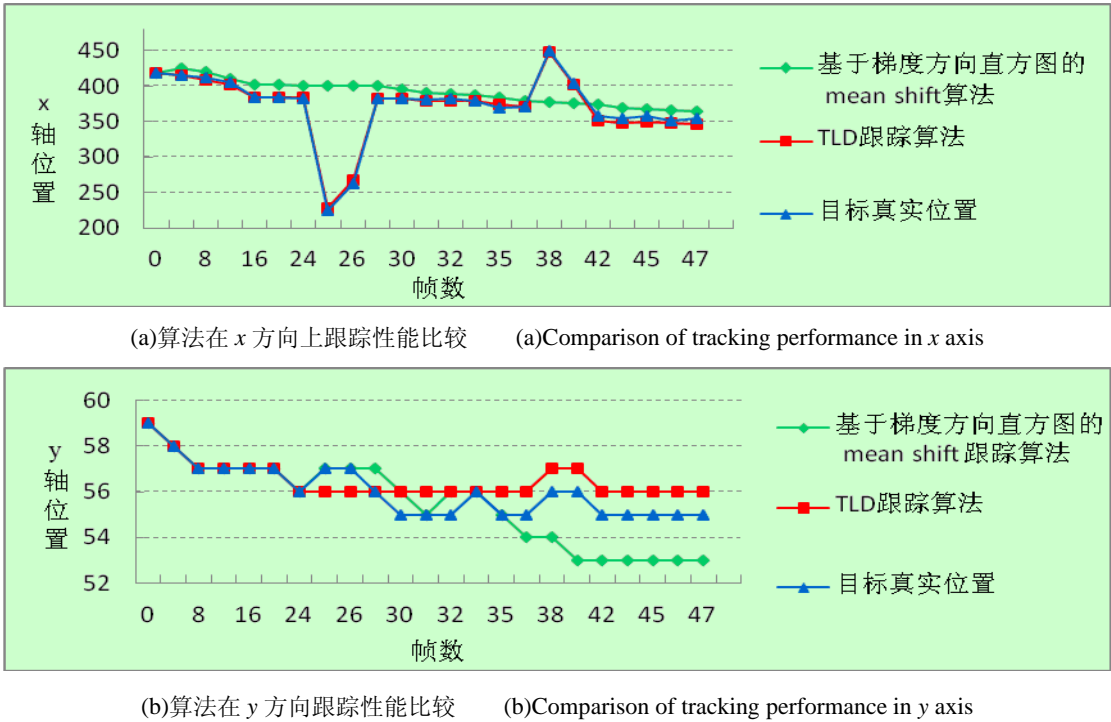


图 11 目标在帧间存在大位移时跟踪性能示意

Fig.11 Diagram of tracking performances when the targets have large displacement in image sequences

3) 目标在某些帧丢失条件下的跟踪仿真实验
此试验模拟目标突然消失的情况(多由于目标被遮挡或相机剧烈运动导致目标移出视场)。为了进

行有效测试,试验中将第 30~37 帧图像进行目标擦除处理以模拟目标丢失。图 12 是此类情况使用 mean shift 跟踪算法得到的效果图,图 13 是使用

TLD 跟踪算法得到的跟踪效果。

在图 12、图 13 中，目标在第 30 帧~第 37 帧发生丢失，但是基于 mean shift 跟踪算法错误地预测了目标位置，造成误跟踪，而 TLD 算法通过随机

全局检测模块对图像中的区域进行检测，同样未获得可疑目标，因此判断目标在当前帧不存在。

图 14 是目标丢失时 2 种算法的跟踪位置曲线图，可以明显地看出基于 mean shift 跟踪算法会在目标丢失期间预测出假的目标位置，而 TLD 算法则在此时不输出目标位置。

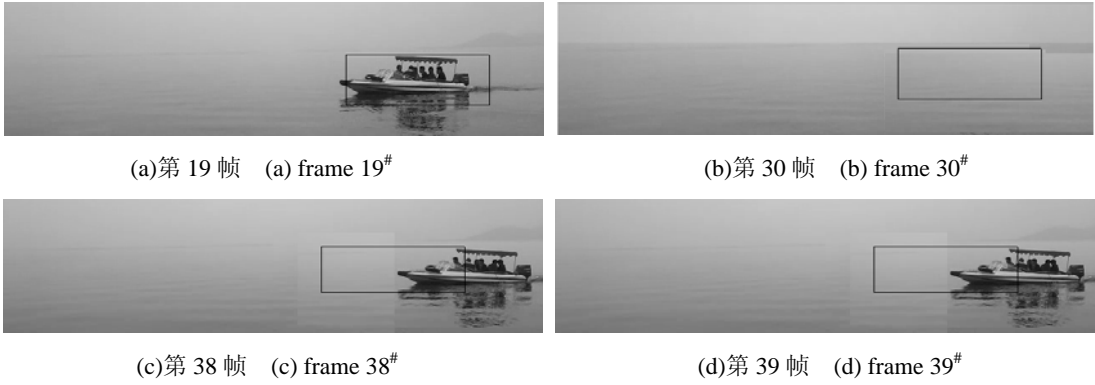


图 12 目标在某些帧丢失时 mean shift 跟踪效果
Fig.12 The tracking results based on mean shift algorithm when the targets loss in some frames

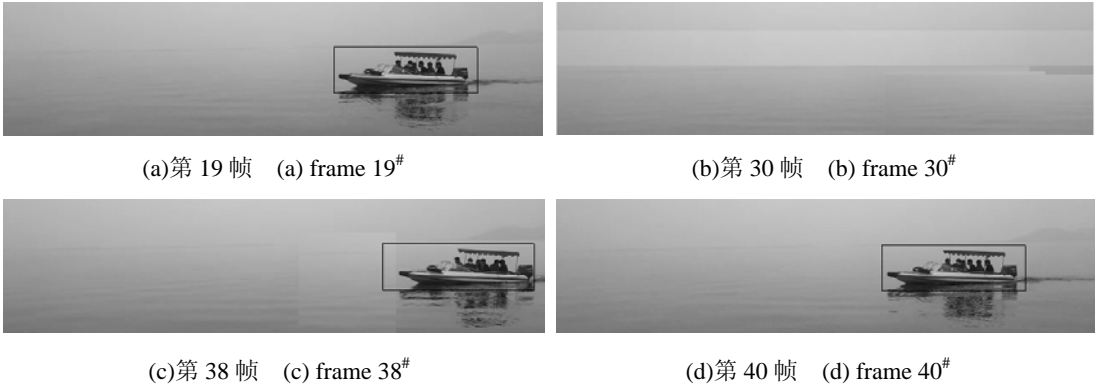


图 13 目标在某些帧丢失时 TLD 跟踪效果
Fig.13 The tracking results based on TLD algorithm when the targets loss in some frames

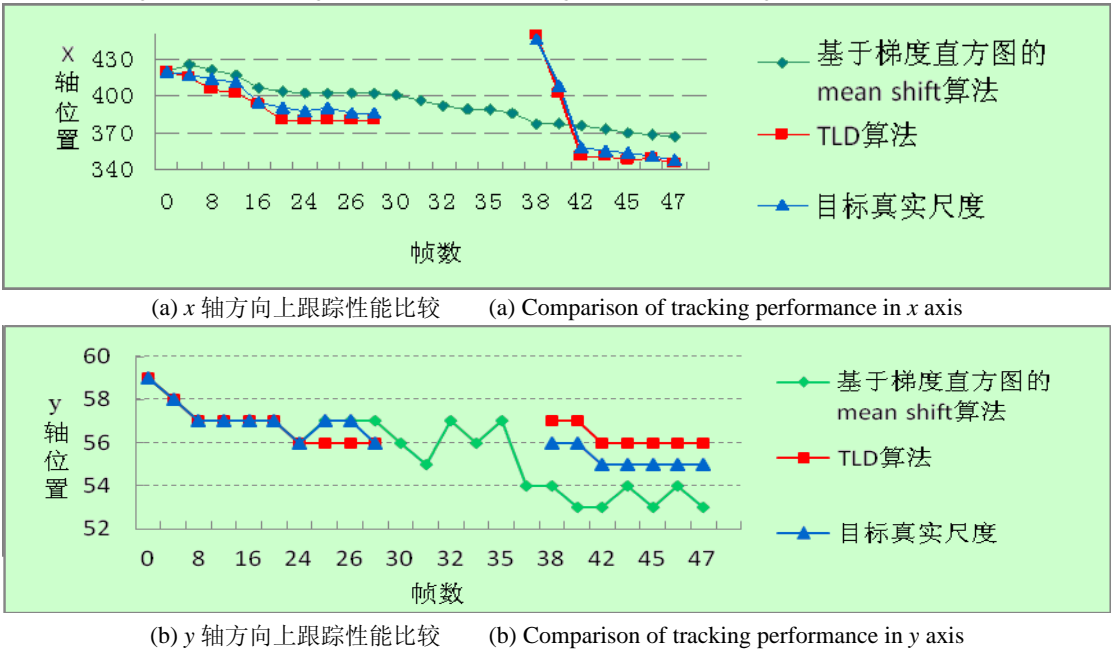


图 14 目标在某些帧丢失时跟踪效果示意 Fig.14 Diagram of tracking performances when the targets loss in some frames

虽然 TLD 算法具有很强的鲁棒性和较高的准确性，但它相对其他传统的跟踪算法耗时也较为巨大，图 15 即是 2 种跟踪算法的时间对比图。

图 15 中可以看出 TLD 算法时间消耗较为巨大，这是因为该算法要对图片进行全局扫描查找相似目标位置以及分类器构建和训练，实际单独的跟

踪模块并不费时间。而且该算法能达到 6~7 fps 的速度，可以满足一般的实时跟踪需要；最重要的是，当跟踪目标发生消失或者不规则运动，其他传统跟踪方案性能不够理想时，TLD 算法可以准确稳定地跟踪到目标，体现了该算法的极强鲁棒性和较高的准确率。

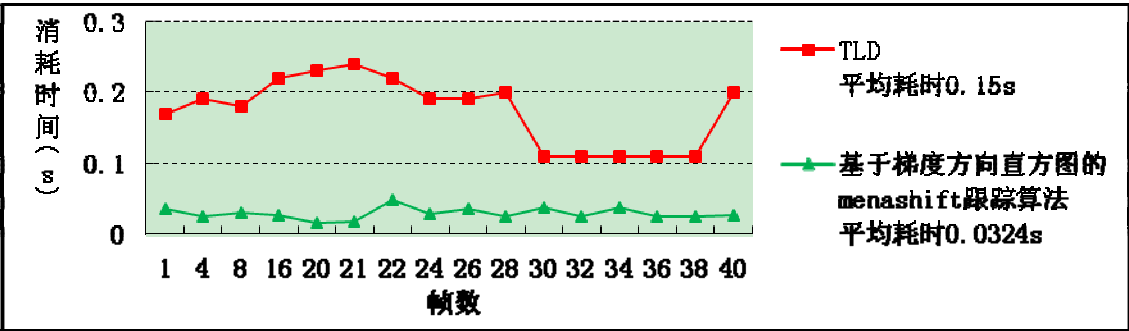


图 15 跟踪算法时间比较

Fig.15 Comparison of time consumption of two tracking algorithms

参考文献：

[1] Huan W, Ming-wu R, Jingyu Y. Object Tracking Based on Genetic Algorithm and Kalman Filter[C]//*International Conference on Computational Intelligence and Security*, 2008: 80-85.

[2] Qin kun X, Xiangjun L, Mina L. Object tracking based on local feature matching[C]//*Fifth International Symposium on Computational Intelligence and Design (ISCID)*, 2012: 399-402 .

[3] Nak Y K, Tae G K. Comparison of Kalman filter and particle filter used for localization of an underwater vehicle[C]//*9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, 2012: 350-352 .

[4] Baohong Y, Dexiang Z, Kui F, et al. Video tracking of human with occlusion based on MeanShift and Kalman filter[C]//*4th Electronic System-Integration Technology Conference (ESTC)*, 2012: 3672-3677.

[5] Jianfang D, Jianxun L, Zhi Z, et al. Face tracking with an Adaptive Adaboost-based Particle Filter[C]//*24th Chinese Control and Decision Conference (CCDC)*, 2012: 3626 - 3631.

[6] Kalal Z, Matas J, Mikolajczyk K. Online learning of robust object detectors during unstable tracking[C]//*12th IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 2009: 1417-1424.

[7] Kalal Z, Mikolajczyk K, Matas J. Face-TLD: Tracking -Learning- Detection applied to faces[C]//*17th IEEE International Conference on Image Processing (ICIP)*, 2010: 3789-3792 .

[8] Ozuysal M, Calonder M, Lepetit V, et al. Fast keypoint recognition using random ferns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(3): 448-461.

[9] Villamizar M, Garrell A, Sanfeliu A, et al. Online human-assisted learning using Random Ferns[C]//*21st International Conference on Pattern Recognition (ICPR)*, 2012: 2821-2824 .

[10] Bosch A, Zisserman A, Muoz X. Image classification using random forests and ferns[C]//*11th IEEE International Conference on Computer Vision*, 2007:1-8 .

[11] Suwon L, Sang-Wook L, Yeong N C, et al. Lightweight Random Ferns using binary representation[C]//*21st International Conference on Pattern Recognition (ICPR)*, 2012: 1342-1345.

[12] Kalal Z, Matas J, Mikolajczyk K. P-N learning: Bootstrapping binary classifiers by structural constraints[C]//*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010: 49-56 .