

# 基于直方图得分和模板更新得分的实时目标跟踪

李静<sup>1,2</sup> 肖春华<sup>3</sup>

<sup>1</sup>(重庆工程学院科技处 重庆 400056)

<sup>2</sup>(重庆市数字影视与新媒体工程技术研究中心 重庆 400056)

<sup>3</sup>(重庆大学计算机学院 重庆 400044)

**摘要** 为了提高目标跟踪对背景颜色和目标形体变化的鲁棒性,在相关滤波器的基础上融合全局颜色直方图,提出一种互补性的跟踪方法。在基于检测的跟踪范式中,建立模板更新和直方图得分的线性组合模型;为了保持相关滤波器的有效性和捕获直方图信息,通过求解两个独立的岭回归问题来学习该模型;根据直方图得分和模板得分的学习结果进行迭代计算,以估计目标下一时刻的位置。实验在 VOT-14 和 OTB-100 两公开视频集中进行,实验结果表明,与其他同类方法相比,该方法在平均覆盖和中心误差方面更优,对颜色、光照和形变均具有明显鲁棒性,且处理速度达到实时标准。

**关键词** 目标跟踪 相关滤波器 直方图 模板更新 鲁棒性 实时

中图分类号 TP391

文献标识码 A

DOI:10.3969/j.issn.1000-386x.2019.02.032

## REAL-TIME TARGET TRACKING BASED ON HISTOGRAM SCORE AND TEMPLATE UPDATING SCORE

Li Jing<sup>1,2</sup> Xiao Chunhua<sup>3</sup>

<sup>1</sup>(Department of Science and Technology, Chongqing Institute of Engineering, Chongqing 400056, China)

<sup>2</sup>(Chongqing Engineering Technology Research Center of Digital Film and Television and New Media, Chongqing 400056, China)

<sup>3</sup>(College of Computer Science, Chongqing University, Chongqing 400044, China)

**Abstract** In order to improve the robustness of target tracking to background color and object shape changes, a complementary tracking method was proposed by fusing the global color histogram based on the related filter. In the detection-based tracking paradigm, we established a linear combination model of template updating and histogram scoring. In order to maintain the validity of related filter and capture histogram information, the proposed model was learned by solving two independent ridge regression problems. The learning results of histogram score and template score were calculated iteratively to estimate the location of the target at the next moment. Experiments were carried out in VOT-14 and OTB-100 public videos. The experimental results show that, compared with other similar methods, the proposed method is better in average coverage and center error. It has obvious robustness to color, illumination and deformation, and the processing speed reaches the real-time standard.

**Keywords** Target tracking Related filter Histogram Template updating Robustness Real-time

## 0 引言

目标跟踪是机器视觉领域热门的基础性研究课题<sup>[1]</sup>。在安防、军事侦察、无人机等众多高新领域应用

广泛。目前,目标跟踪的难点和挑战是如何较好地解决目标的形态变化、不准确预测造成的漂移现象、低分辨率的目标识别、多目标的遮挡和身份切换等问题<sup>[2-3]</sup>。

在已有的研究成果中,跟踪算法按照不同策略,可

以分为基于学习的算法、基于滤波的算法和多跟踪器模式。基于学习的算法是一种自适应方法,即在线学习模式。如 Struck<sup>[4]</sup>是一个性能较好、结构简单的在线目标检测与跟踪算法,该算法旨在将目标定位的结构化输出最小化,但较高的计算量限制了特征和训练样本的数量。文献[5]提出基于 HOG 在线多实例学习的目标跟踪算法,该算法采用 HOG 特征值提取方式,结合在线多实例学习技术,其适合于目标平移、旋转和遮挡等情况下的跟踪。

TLD(Tracking Learning Detection)算法<sup>[6]</sup>采用 NP 学习方法对训练出现的正负样本进行分类和纠错,通过修改训练策略提高预测的准确性,减少了目标漂移的可能性。文献[7]将协同训练应用到跟踪中,对两个独立的、使用了不同特征的支持向量机进行学习,随后从合并的分数中得出硬性否定。

基于滤波的跟踪大多会局限于对一个刚性模板的学习<sup>[8]</sup>。当目标发生形状变化时,跟踪会变得很困难,解决思路是采用一个对形状变化不敏感的特征形式,如 KCF 方法<sup>[9]</sup>。一般可以将直方图与相关滤波器正交,但仅靠直方图通常无法有效表征目标,为此,文献[10]对来自每个像素的投票进行累加,使用投票后得分最高的位置像素来估计目标的范围。但是,这类方法并不具有普遍适用性,假设条件较多。

对于一些特殊形式的运动方式,需要对跟踪形式进行改进。例如针对低秩假设的粒子滤波跟踪器难以跟踪突然运动目标的局限性,文献[11]提出时间受限的反向低秩视觉跟踪算法,利用反向低秩模型表示目标模板和背景模板,并利用混合范数进行局部约束,以保证目标外观的局部一致性,以及相邻帧之间的突然变化。

为了降低预测不准确性,一般可以结合多个估计手段,使跟踪器的弱点得到补偿。如文献[12]提出一种低复杂度交互式多模型滤波器,该跟踪算法组合了一些交互式滤波器,并且利用了整个滤波器组的卡尔曼滤波增益矩阵。文献[13]使用隐马尔科夫模型 HMM(Hidden Markov Model)对目标轨迹以及不同时刻的跟踪器可靠性进行建模。但这类方法缺点是各跟踪器的复杂度和兼容性需要统筹考虑。

本文主要创新之处是所提跟踪器结合了两方面的得分:目标模板得分和直方图得分,建立线性组合模型。这两方面具备一些重要属性,即:得分在量值上具有相似性,可靠性较高,且目标的位置预测由置信度更高的一方所主导。这样使得跟踪器对互补性要素具有敏感性,对局部变化不敏感,所学习的模型对颜色变化和形变均具有一定的鲁棒性。另外,为保持实时的处

理速度,本文还利用了每个图像块表示内在结构,并将其作为两个独立的岭回归框架<sup>[14]</sup>进行求解。实验结果验证了所提方法的优越性。

## 1 模型说明

本文目标跟踪采用了基于检测的跟踪范式,即:在帧  $t$  中,从集合  $S_t$  选出目标在图像  $x_t$  中矩形跟踪框的位置得分  $p_t$ :

$$p_t = \arg \max_{p \in S_t} f(T(x_t, p); \theta_{t-1}) \quad (1)$$

式中:函数  $T(\cdot)$  表示图像变换,  $f(T(x, p); \theta)$  表示参数  $\theta$  为图像  $x$  中的矩形跟踪框  $p$  分配的一个得分函数。为了跟踪的需要,应该选择合适的参数将损失函数  $L(\theta; \mathcal{X}_t)$  最小化,该损失函数取决于以往图像目标在这些图像中的位置  $\mathcal{X}_t = \{(x_i, p_i)\}_{i=1}^t$ , 该位置的  $\theta_t$  定义如下:

$$\theta_t = \arg \min_{\theta \in Q} \{L(\theta; \mathcal{X}_t) + \lambda R(\theta)\} \quad (2)$$

式中:  $Q$  为参数  $\theta$  的空间表示, 本文使用正则化项  $R(\theta)$  限制模型的复杂度,以防止过拟合。

式(1)中得分函数  $f(x)$  定义为模板(tmp)和直方图(hist)得分的一个线性组合:

$$f(x) = \gamma_{\text{tmp}} f_{\text{tmp}}(x) + \gamma_{\text{hist}} f_{\text{hist}}(x) \quad (3)$$

为了在滑动窗口的搜索中,高效计算得分函数,窗口图像特征共享重叠窗口,模板得分使用卷积快速计算,直方图得分通过单积分图计算获得。

在理想情况下,每帧图像训练损失函数的形式为:

$$\ell(x, p, \theta) = d(p, \arg \max_{q \in S} f(T(x, q); \theta)) \quad (4)$$

式中:  $d(p, q)$  定义为:当正确矩形框为  $p$  时选择矩形框  $q$  的代价。虽然该函数是一个非凸性函数,但可使用结构化输出学习对目标的界限进行优化。然而,该优化的缺点在于计算成本很高,限制了可用特征和训练样本的数量。相比之下,相关滤波器采用了简单的最小二乘原理,可通过将特征图像的循环位移作为样本,使用较高维特征表示,从一个相对数量较大的训练样本中学习。该方法在跟踪中效果较好,同时保持了较高的运行速率。

因此,为了保持相关滤波器的速度和有效性,同时不忽略直方图得分所捕获到的信息,本文通过求解两个独立的岭回归问题学习所提模型,即:

$$\begin{aligned} h_t &= \arg \min_h \left\{ L_{\text{tmp}}(h; \mathcal{X}_t) + \frac{1}{2} \lambda_{\text{tmp}} \|h\|^2 \right\} \\ \beta_t &= \arg \min_{\beta} \left\{ L_{\text{hist}}(\beta; \mathcal{X}_t) + \frac{1}{2} \lambda_{\text{hist}} \|\beta\|^2 \right\} \end{aligned} \quad (5)$$

式中:参数  $h$  可以使用相关性滤波公式快速得出。最后,选取两个模型得分的凸组合,设  $\gamma_{\text{tmp}} = 1 - \alpha$ ,  $\gamma_{\text{hist}} =$

$\alpha$ , 其中因子  $\alpha$  是在验证集上选择的参数。以上两个模型得分函数的参数都将得到优化, 目标的分数分配为 1, 其他窗口的分数分配为 0。所提模型的总体示意图如图 1 所示。

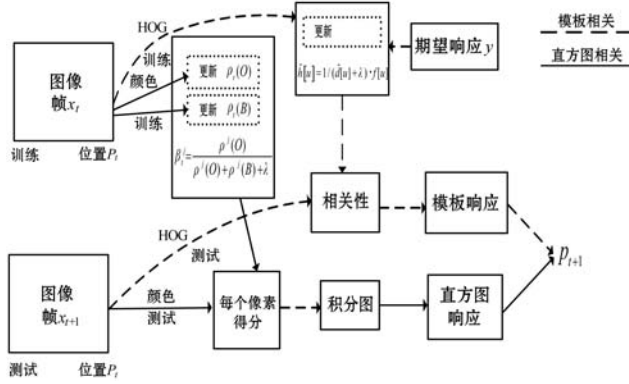


图 1 所提模型的总体示意图

## 2 优化与学习过程

### 2.1 在线最小二乘优化

本文采用最小二乘和二次正则算子有两个好处:

1) 在闭合形式中得到解; 2) 内存要求不会随样本数量的增加而增加。如果  $L(\theta; \mathcal{X})$  是得分  $f(x; \theta)$  的一个凸二次函数, 而  $f(x; \theta)$  在模型参数  $\theta$  中是线性的 (以保持凸性), 那么则会存在着一个矩阵  $A_t$  和一个向量  $b_t$ , 使得:

$$L(\theta; \mathcal{X}) + \lambda \|\theta\|^2 = \frac{1}{2} \theta^T (A_t + \lambda I) \theta + b_t^T \theta \quad (6)$$

而无论  $\mathcal{X}_t$  的大小如何, 都足以确定解  $\theta_t = (A_t + \lambda I)^{-1} b_t$ 。

如果采用一个递归定义的损失函数:

$$L(\theta; \mathcal{X}) = (1 - \eta) L(\theta; \mathcal{X}_{t-1}) + \eta \ell(x_t, p_t, \theta) \quad (7)$$

式中:  $\eta$  为自适应率, 则可维持:

$$A_t = (1 - \eta) A_{t-1} + \eta A'_t \quad (8)$$

$$b_t = (1 - \eta) b_{t-1} + \eta b'_t$$

式中:  $A'_t$  和  $b'_t$  通过下式定义每帧图像的损失。

$$\ell(x_t, p_t, \theta) = \frac{1}{2} \theta^T A'_t \theta + \theta^T b'_t \quad (9)$$

式中:  $A_t$  表示从第 1 帧到第  $t$  帧估计出的参数, 而  $A'_t$  则表示仅从第  $t$  帧估计出的参数。这些参数足以得到一个解, 且在特征的数量较小或矩阵有冗余 (例如稀疏矩阵或低秩矩阵) 的情况下, 计算和存储的成本通常相对较小。

### 2.2 对模板得分进行学习

在最小二乘相关滤波中, 每帧图像损失为:

$$\ell_{\text{tmp1}}(x, p, h) = \left\| \sum_{k=1}^K h^k \star \mathcal{O}^k - y \right\|^2 \quad (10)$$

式中:  $h^k$  表示多通道图像  $h$  的通道  $k$ ,  $\mathcal{O}$  为  $\mathcal{O}_{T(x,p)}$  的缩写,  $y$  为期望, 而  $\star$  表示周期性互相关。对模板得分进行学习的正则化目标函数  $\ell_{\text{tmp1}}(x, p, h) + \lambda \|h\|^2$  的极小值表示为:

$$\hat{h}[u] = (\hat{s}[u] + \lambda I)^{-1} \hat{r}[u] \quad (11)$$

对于所有的  $u \in T$ ,  $\hat{s}[u]$  是一个包含矩阵元素  $\hat{s}^{ij}[u]$ , 大小为  $K \times K$  的矩阵,  $\hat{r}[u]$  则是包含元素  $\hat{r}^i[u]$  的一个  $K$  维向量。

在实践中, 本文没有对式 (11) 进行计算, 而是采用了尺度估计的相关滤波器<sup>[15]</sup>进行近似。

模板在线更新表示为:

$$\begin{aligned} \hat{d}'_t &= (1 - \eta_{\text{tmp1}}) \hat{d}'_{t-1} + \eta_{\text{tmp1}} \hat{d}'_t \\ \hat{r}'_t &= (1 - \eta_{\text{tmp1}}) \hat{r}'_{t-1} + \eta_{\text{tmp1}} \hat{r}'_t \end{aligned} \quad (12)$$

式中:  $\hat{d}'$  和  $\hat{r}'$  分别为式 (11) 尺度估计的相关滤波器的参数。

### 2.3 对直方图得分进行学习

理想情况下, 直方图得分应该学习来自于每个图像的一个样本集合, 其中包括作为一个正例的正确位置。设  $W$  表示矩形窗  $q$  和与之相对应的回归目标  $y \in \mathbb{R}$  的一个成对集合  $(q, y)$ , 包括正例  $(p, 1)$ 。则每帧图像的损失为:

$$\ell_{\text{hist}}(x, p, \beta) = \sum_{(q,y) \in W} (\beta^T [\sum_{u \in H} \psi_{T(x,q)}[u]])^2 \quad (13)$$

通过对一个  $M \times M$  的方程组进行求解, 得出  $M$ -通道特征变换  $\psi$  的解, 求解该方程组的内存需求为  $O(M^2)$ , 时间需求为  $O(M^3)$ 。如果特征的数量较大, 则上述方法不可行。虽然存在着其他的迭代方法对矩阵进行分解<sup>[10]</sup>, 例如坐标下降法、共轭梯度和双坐标下降, 但是仅靠这些方法依然难以实现较高的帧速率。

本文提出了形式的特征为  $\psi[u] = e_{k[u]}$ , 式中的向量  $e$  在下标为  $i$  时等于 1, 其他情况下则等于 0。考虑的特征为量化的 RGB 颜色特征, 直方图得分可被视为一个平均投票。因此, 为了提高效率, 本文建议将线性回归独立的应用到目标  $\mathcal{O}$  和背景区域  $\mathcal{B} \in \mathbb{Z}^2$  的每个特征像素, 使用每帧图像目标:

$$\begin{aligned} \ell_{\text{hist}}(x, p, \beta) &= \frac{1}{|\mathcal{O}|} \sum_{u \in \mathcal{O}} (\beta^T \psi[u] - 1)^2 + \\ &\quad \frac{1}{|\mathcal{B}|} \sum_{u \in \mathcal{B}} (\beta^T \psi[u])^2 \end{aligned} \quad (14)$$

式中:  $\psi$  为  $\psi_{T(x,p)}$  的简写。引入 one-hot 假设, 目标分解为每特征维数独立项:

$$\ell_{\text{hist}}(x, p, \beta) = \sum_{j=1}^M \left[ \frac{N^j(\mathcal{O})}{|\mathcal{O}|} \cdot (\beta^j - 1)^2 + \right]$$

$$\frac{N^j(\mathcal{B})}{|\mathcal{B}|} \cdot (\beta^j)^2 \quad (15)$$

式中:  $N^j(\mathcal{A}) = |\{u \in \mathcal{A} : k[u] = j\}|$  是  $\mathcal{O}_{T(x,p)}$  的区域  $\mathcal{A}$  中的像素数量, 特征  $j$  为非零  $k[u] = j$ 。则相关联的岭回归问题的解为:

$$\beta_t^j = \frac{p^j(\mathcal{O})}{p^j(\mathcal{O}) + p^j(\mathcal{B}) + \lambda} \quad (16)$$

对于每个特征维数  $j = 1, 2, \dots, M$ , 其中  $\rho^j(\mathcal{A}) = N^j(\mathcal{A})/|\mathcal{A}|$  为一个区域中  $j$  为非零的像素比例。本文对模型参数进行更新:

$$\rho_t(\mathcal{O}) = (1 - \eta_{\text{hist}})\rho_{t-1}(\mathcal{O}) + \eta_{\text{hist}}\rho'_t(\mathcal{O}) \quad (17)$$

$$\rho_t(\mathcal{B}) = (1 - \eta_{\text{hist}})\rho_{t-1}(\mathcal{B}) + \eta_{\text{hist}}\rho'_t(\mathcal{B})$$

式中:  $\rho_t(\mathcal{A})$  是  $\rho_t^j(\mathcal{A})$  的向量,  $j = 1, 2, \dots, M$ 。

### 3 实验与分析

本文在视频集 VOT-14<sup>[2]</sup> 和 OTB-100<sup>[16]</sup> 上对所提跟踪器和其他方法进行了比较, 实验中所使用的重要参数如表 1 所示。

表 1 本文实验中所采用的参数列表

模板学习率 $\eta_{\text{tmp}}$	0.01
直方图学习率 $\eta_{\text{hist}}$	0.04
颜色特征	RGB
颜色直方图	32 × 32 × 32
合并参数 $\alpha$	0.3
固定区域	150 <sup>2</sup>
HOG 单元格尺寸	4 × 4

#### 3.1 VOT-14 视频集

从 VOT-14 视频集的 394 个视频中选择 25 个序列, 以代表相机运动、遮挡、光照变化、尺寸和运动, 并在选出的序列上对跟踪器进行性能比较, 该部分列举了有代表性 8 个视频的跟踪结果。比较的跟踪有文献[4]提出的 Struck 跟踪器, 文献[5]提出的多实例学习的 HOG 跟踪方法, 经典的 TLD 跟踪器<sup>[6]</sup>, 以及改进的 KCF 跟踪器<sup>[9]</sup>。

使用的评价指标如下所示:

1) 跟踪器在一个序列上的准确度使用交集与并集之比:  $S_t = \frac{|r_t \cap r_{GT}|}{|r_t \cup r_{GT}|}$ , 表示该跟踪器预测的包围框  $r_t$

和实际包围框  $r_{GT}$  之间的平均每帧重叠。

2) 包围框中心误差可以量化跟踪器目标预测中心

$TT_k^c$  与实际包围框中心  $GT_k^c$  的误差:  $\sqrt{\frac{1}{N} \sum_{k=1}^N \|TT_k^c - GT_k^c\|_2^2}$ ,

其中,  $N$  是视频总帧数。

3) 跟踪器的鲁棒性为其在视频中失败的次数。当  $S_t$  变成零时, 则判定出现了一次跟踪失败。由于该基准着眼于短时间跟踪, 一个跟踪器在出现跟踪失败后, 会在 5 帧之后被自动重新初始化到地面位置。

各跟踪器对各视频序列的平均每帧重叠结果如表 2 所示, 平均重叠值越大, 其对应的跟踪器越优, 每一行的最优结果加粗显示。可以看出本文方法在大多数情况下, 其重叠均达到最大。即使在次优情况下, 也与最优相差不多。在地铁视频中表现不佳, 可能是由于矩形包围框的中心位置造成的。不同跟踪器对包围框中心误差的比较如表 3 所示。可以看出, 本文方法的中心误差最低, 跟踪的中心位置把握得更准确。25 个视频的综合评价结果如表 4 所示。可以看出, 本文方法的平均重叠更大, 跟踪失败次数更少, 综合评价得分最高。

表 2 各跟踪器对视频序列的重叠值

视频序列	Struck	多实例 HOG	TLD	改进 KCF	本文方法
篮球	0.355	0.109	0.132	0.101	0.427
运动员	0.307	0.202	0.188	0.290	0.426
车辆	0.543	0.516	0.459	0.271	0.620
David	0.586	0.293	0.317	0.476	0.607
地铁视频	0.089	0.301	0.157	0.195	0.291
熊猫	0.273	0.452	0.410	0.510	0.572
自行车	0.172	0.288	0.151	0.178	0.290
无人机	0.186	0.189	0.153	0.248	0.295

表 3 各跟踪器对视频序列的中心误差

视频序列	Struck	多实例 HOG	TLD	改进 KCF	本文方法
篮球	101	106	131	84	11
运动员	41	127	81	9	6
车辆	101	121	129	45	29
David	18	19	70	26	17
地铁视频	140	38	110	44	27
熊猫	78	38	25	37	17
自行车	110	16	192	18	10
无人机	60	6	112	28	9

表 4 25 个视频序列中的综合排名结果

跟踪器	平均重叠	平均失败次数
本文方法	0.536	81
Struck	0.461	152
多实例 HOG	0.492	124
TLD	0.376	205
改进 KCF	0.471	163

为了研究形变对跟踪器的影响,选择一些形体不停发生变换的视频,如做复杂动作的运动员视频,如图2所示。其中第一行是多实例 HOG 方法的跟踪结果,第二行是改进 KCF 方法的跟踪结果,第三行是本文方法的结果。可以看出,本文方法在运动员的运动过程均得到了良好的跟踪。多实例 HOG 方法和 KCF 方法受到背景颜色的严重干扰,同时由于形变造成模板特征发生了较大变化,而本文方法充分利用目标模板和直方图得分的互补性对形变和背景具有较好鲁棒性。

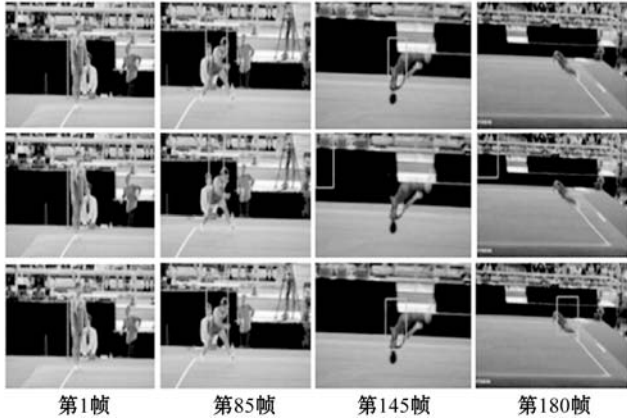


图2 运动员视频中外观变化实验

为了研究光照变化对跟踪器的影响,选取了 David 视频中明暗变化的序列部分,该部分伴有较小形变过程,如图3所示。比较的方法与图2外观变化实验相同。可以看出多实例 HOG 和改进 KCF 方法由于更新误差积累导致了目标漂移,而本文方法对光照变化具有较好的鲁棒性。一般来说,基于相关滤波的跟踪器在应对运动、光照和尺度等变化时具有一定的鲁棒性,但局限于刚性模板的学习,当目标在一个视频序列中发生形变时,该类跟踪器会表现不佳。在所用特征中,颜色直方图会丢弃每个像素的位置信息,对形状变化具备鲁棒性,但不能很好地应对光照变化,且当颜色辨识度不高时通常无法有效地从背景中辨识。鉴于直方图和模板所具备的互补性,本文结合两个图像分块表示,对形状和颜色均具备鲁棒性。

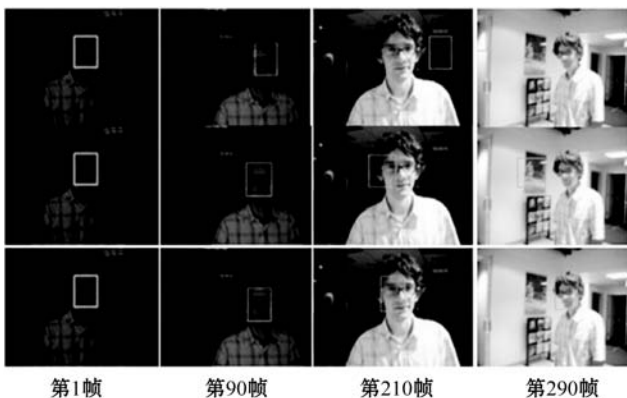


图3 David 视频序列中光照变化实验

### 3.2 OTB-100 视频集

与 VOT-14 基准一样,OTB-100<sup>[16]</sup>的理念是在准确度和失败的鲁棒性两方面对跟踪器的性能进行评价,OTB-100 视频集已经成为目标跟踪数据集的一个标准。准确度通过预测跟踪器的矩形包围框和实际情况的交集与并集之比进行度量。但用到了一个阈值  $t_0$ ,当该比值高于阈值  $t_0$  时,则检测到一次成功的跟踪。为了不将该阈值设为一个特定数值,在成功率曲线下方处于不同数值的  $t_0$  区域作为一次最后得分,采用空间鲁棒性评价和时序鲁棒性评价评估跟踪器<sup>[16]</sup>。

本文使用了与 VOT-14 中相同的代码和参数进行测试,结果如图4和图5所示,可以看出,本文方法在成功率方面更优,总体上优于其他方法。其中,多实例学习的 HOG 整体比较差,鉴于该方法使用相关滤波器学习一个颜色模板,因此,本文方法的优越性表明,通过结合模板和直方图得分所实现的性能改善,不能单一地归功于对颜色的引入。

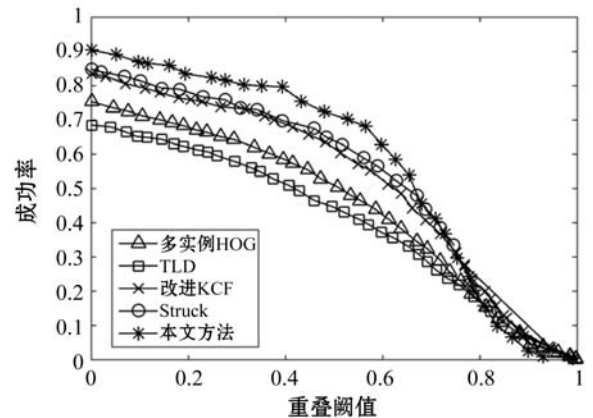


图4 时序鲁棒性评价

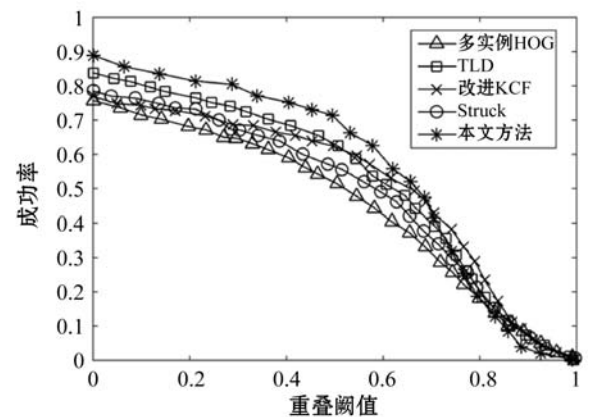


图5 空间鲁棒性评价

### 3.3 效率

本文采用 MATLAB 和 C 混合编程,在配置了 Intel Core i5-4790K @ 8.0 GHz 的台式计算机上以每秒大约 80 帧的速率运行。但是,通过对模型计算出的分块尺

寸进行调节,可以牺牲较小的性能,以实现一个较高的帧速率。例如使用 HOG 的单元格尺寸为  $2 \times 2$ ,失败次数的增加较少,而速率则超过了每秒 100 帧。

### 3.4 学习率实验

本文分别将学习率  $\eta_{\text{templ}}$  和  $\eta_{\text{hist}}$  用于模板和直方图模型更新,以确定从之前帧中得到的旧证据替换为当前帧中得到的新证据。一般情况下,学习率越低,则从先前帧中学习到的模型实例相关性越高。实验结果表明, $\eta_{\text{templ}}$  和  $\eta_{\text{hist}}$  均大约为 0.01,可以达到最大程度的鲁棒性。

### 3.5 因子 $\alpha$ 的影响

本文方法的准确度(平均重叠值)受因子  $\alpha$  的影响较大,其具体情形如图 6 所示,而式(3)中的  $\gamma_{\text{templ}}$  和  $\gamma_{\text{hist}}$  由  $\alpha$  控制:在  $\alpha = 0.3$  附近,本文方法的性能最优。鲁棒性遵循类似的趋势。另外,图 6 还说明了对两个岭回归问题的密集响应进行合并的策略所达到的性能显著优于仅对最终估计进行插值的策略,即:具有相容且互补的密集响应模型能够获得较优的结果。

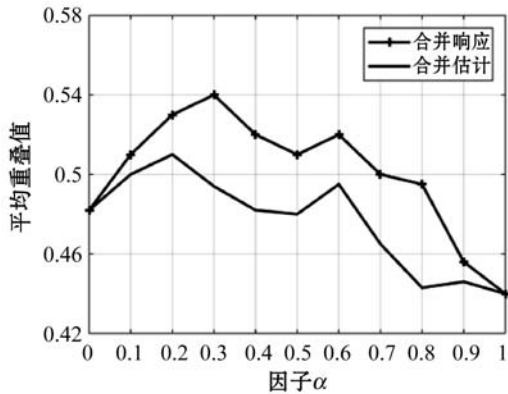


图 6 平均重叠值与因子  $\alpha$  的关系

## 4 结 语

本文在岭回归框架下,对模板和直方图得分进行独立学习,并将两者进行线性组合,以实现在实时跟踪的基础上对颜色、形变和光照变化的鲁棒性。实验结果表明提出的跟踪器性能优于其他跟踪器,且计算量需求较低。

未来将考虑多个独立跟踪器进行多目标跟踪,同时也会考虑多目标跟踪中的挑战性问题,如身份切换和目标函数收敛等问题。

## 参 考 文 献

[1] 李静,王军政. 图像检测与目标跟踪技术[M]. 北京:北

京理工大学出版社, 2013.

- [2] Kristan M, Pflugfelder R, Leonardis A, et al. The visual object tracking VOT2014 challenge results [C]//Computer Vision—ECCV 2014 Workshops. Springer International Publishing, 2014: 191–217.
- [3] 王保宪. 复杂背景下的视频目标跟踪算法研究[D]. 北京:北京理工大学, 2016.
- [4] Hare S, Golodetz S, Saffari A, et al. Struck: structured output tracking with kernels[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 23(5): 263–270.
- [5] 刘哲, 陈息, 郑紫微. 基于 HOG 与多实例在线学习的目标跟踪算法[J]. 计算机工程, 2015, 41(1): 158–163.
- [6] Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1409–1422.
- [7] 胡昭华, 徐玉伟, 赵孝磊, 等. 基于支持向量机的多特征选择目标跟踪[J]. 应用科学学报, 2015, 33(5): 502–517.
- [8] 李介谷. 计算机视觉的理论和实践[M]. 上海:上海交通大学出版社, 1991.
- [9] 郑武兴, 王春平, 付强. 改进的 KCF 红外空中目标跟踪方法[J]. 激光与红外, 2017, 47(12): 1553–1558.
- [10] Duffner S, Garcia C. PixelTrack: A fast adaptive algorithm for tracking non-rigid objects[C]//IEEE International Conference on Computer Vision. IEEE Computer Society, 2013: 2480–2487.
- [11] Yang Y, Hu W, Xie Y, et al. Temporal restricted visual tracking via reverse-low-rank sparse learning [J]. IEEE Transactions on Cybernetics, 2017, 47(2): 485–498.
- [12] Khalid S S, Abrar S. A low-complexity interacting multiple model filter for maneuvering target tracking [J]. AEUE—International Journal of Electronics and Communications, 2017, 73(12): 157–164.
- [13] 顾苏杭, 戎海龙, 马正华. 融合 SIFT 特征和 HMM 的运动目标识别与跟踪算法[J]. 科技通报, 2017, 33(11): 210–215.
- [14] 王之毅, 毕彥彦, 熊磊, 等. 基于稀疏编码和岭回归的文本图像复原算法[J]. 计算机应用, 2017, 37(9): 2648–2651.
- [15] Danelljan M, Hager G, Khan F S, et al. Learning spatially regularized correlation filters for visual tracking[C]//IEEE International Conference on Computer Vision. IEEE Computer Society, 2015: 4310–4318.
- [16] Wu Y, Lim J, Yang M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834–1848.