

文章编号 1004-924X(2018)08-2112-10

## 利用视觉显著性和扰动模型的上下文感知跟踪

张 博<sup>1\*</sup>, 江沸菠<sup>2</sup>, 刘 刚<sup>1,3</sup>

- (1. 长沙师范学院 信息与工程系, 湖南 长沙 410100;
2. 湖南师范大学 物理与信息科学学院, 湖南 长沙 410081;
3. 中南大学 物理与电子学院, 湖南 长沙 410083)

**摘要:**为了解决背景嘈杂、遮挡、形变和尺度变化情况下目标跟踪问题,提出利用视觉显著性和扰动模型的上下文感知跟踪。本文以相关滤波算法为基础,将目标周围的上下文信息引入到分类器学习过程中,构造了上下文感知相关跟踪,提高了算法鲁棒性;同时引入直方图扰动模型,利用加权融合的方法获得目标响应图,以此估计目标位置变化;最后利用视觉显著性构建目标稀疏显著性图,解决严重遮挡情况下的目标重定位问题,并利用尺度估计策略解决目标尺度变化问题。利用公开数据集测试算法性能,并与 8 种流行跟踪算法进行比较。实验结果表明,本文算法的跟踪精确度得分和成功率得分分别为 0.695 和 0.708,均优于其它算法。与传统的相关滤波算法相比,所提算法能很好地解决背景嘈杂、遮挡、形变和尺度变化等复杂下的目标跟踪问题,具有一定理论研究价值和工程实用价值。

**关键词:**目标跟踪;上下文感知;扰动模型;视觉显著性;相关滤波

**中图分类号:**TH703 **文献标识码:**A **doi:**10.3788/OPE.20182608.2112

## Context-aware tracking based on a visual saliency and perturbation model

ZHANG Bo<sup>1\*</sup>, JIANG Fei-bo<sup>2</sup>, LIU Gang<sup>1,3</sup>

- (1. Department of information and Engineering, Changsha Normal University, Changsha 410100, China;
2. College of Physics and Information Science, Hunan Normal University, Changsha 410081, China;
3. Physical Science and Electronics, Central South University, Changsha 410083, China)

\* Corresponding author, E-mail: zb801121@126.com

**Abstract:** To solve the problem of target tracking in the presence of background noise, occlusion, deformation and scale variation, a context-aware tracking algorithm based on a visual saliency and perturbation model was proposed. First, the proposed algorithm was based on the correlation filtering algorithm. The contextual information of the target was introduced into the classifier learning process. The context-aware correlation filter was then constructed, which improves the robustness of the algorithm. Meanwhile, the histogram perturbation model was introduced. The target response map was calculated using the weighted fusion method to estimate the target position change. Finally, the target saliency map was constructed using visual saliency to solve the target relocation problem under occlu-

收稿日期:2017-11-24;修订日期:2018-01-17.

基金项目:国家自然科学基金青年科学基金资助项目(No. 41604117);院级重点项目资助(No. XYZD2016090)

sion problem. The scale estimation strategy was used to solve the problem of target scale variation. The algorithm performance was tested using open-source datasets and was compared with eight popular tracking algorithms. The experimental results demonstrate that the accuracy and success rate of the algorithm are 0.695 and 0.708, respectively, which are better than other algorithms. Compared with the traditional correlation filtering algorithm, the proposed algorithm can solve the target tracking problem with complex background noise, occlusion, deformation and scale changes. It has a certain theoretical research value and practical value of engineering.

**Key words:** target tracking; context-aware; perturbation model; visual saliency; correlation filter

## 1 引言

视觉目标跟踪作为计算机视觉领域中的一个重要研究分支,其已经被广泛地应用到无人驾驶、人机交互、智能监控和虚拟现实等诸多领域。在实际跟踪环境中,跟踪目标往往受背景嘈杂、尺度变化、形变和遮挡等因素影响,如何解决复杂背景下的精确目标跟踪问题仍是视觉目标跟踪领域中的一个难点与热点。

视觉目标跟踪算法主要可以分为判别式跟踪<sup>[1-4]</sup>和生成式跟踪<sup>[5-6]</sup>。生成式跟踪需要构建复杂的目标外观模型,并以粒子滤波算法为框架,其算法计算复杂度较高,往往不能实现在线实时跟踪;而生成式方法不需要构建复杂目标外观模型,其将跟踪看做是在线分类问题,将跟踪问题转化为求解最优分类器的过程,利用分类器将跟踪目标从背景中区分出来。

近几年,判别式跟踪方法得到快速发展,涌现出大量优秀的跟踪算法。Bolme等<sup>[7]</sup>提出最小均方误差下输出平方和(Minimum Output Sum of Squared Error, MOSSE)的相关滤波跟踪,首次将相关滤波理论引入到目标跟踪中,通过在灰度图像上训练得到最小均方误差下的相关滤波分类器,将时域内的计算转换到频域,极大地降低了计算花费并提高了目标跟踪的速度;Henriques等<sup>[8]</sup>提出核相关滤波(Kernelized Correlation Filter, KCF)跟踪算法,用梯度方向直方图(Histogram of Oriented Gradient, HOG)特征表征目标的外观模型,利用核策略和循环结构将时域内的卷积操作转换为频域内的点积操作,显著提高了核相关滤波算法的跟踪精确度和跟踪鲁棒性;Zhang等<sup>[9]</sup>提出时空上下文(Spatio-Temporal context, STC)跟踪算法,通过贝叶斯框架对目标

与其周围的局部上下文信息进行时空关系建模,得到目标与其周围背景信息的统计相关性,利用目标置信度图估计目标所在位置;Ma等<sup>[10]</sup>提出一种长期相关跟踪(Long-term Correlation Tracking, LCT)算法,引入在线重检测机制,利用随机森林分类器解决目标跟踪失败情况下的重定位问题,提高了算法鲁棒性。

但是上述方法未能很好解决背景嘈杂、形变和遮挡情况下的目标跟踪问题,本文提出利用视觉显著性和扰动模型的上下文感知跟踪。在传统的相关滤波跟踪算法中因为利用局部特征表征目标外观模型,其对目标形变与嘈杂背景下的干扰较为敏感,且无法解决遮挡情况下的目标模板更新问题和重定位问题,所以本文通过加权响应的方式将上下文感知相关滤波器与直方图扰动模型进行融合,显著提高嘈杂背景与目标形变下分类器的分类能力;利用视觉显著性图引入重检测机制,解决严重遮挡下的目标重定位问题;利用自适应模板更新策略,解决复杂背景下的目标遮挡问题;最后利用公开数据<sup>[11]</sup>测试本文算法。

## 2 相关跟踪

相关滤波算法<sup>[8]</sup>中将目标跟踪问题转化为求解最小均方误差下的岭回归问题,以此获得最优分类器参数,其定义为:

$$\epsilon = \operatorname{argmin}_{\omega} \sum_i \|x_i \omega - y_i\|_2^2 + \lambda \|\omega\|_2^2, \quad (1)$$

式中: $\lambda$ 是正则化因子,作用是防止分类器过拟合, $\omega$ 是分类器参数, $x_i$ 表示由基础样本循环偏移产生的训练样本, $y_i$ 为训练样本 $x_i$ 对应的训练标签或者期望输出标签,标签函数 $y$ 服从高斯分布且 $y_i \in [0, 1]$ 。

求解公式(1)的闭环解,获得最小均方误差下

的最优分类器为:

$$\omega = \frac{\mathbf{X}^T \mathbf{y}}{\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I}}, \quad (2)$$

式中:  $\mathbf{X}$  表示由循环偏移样本组成的循环数据矩阵, 上标 T 表示矩阵的转置,  $\mathbf{I}$  表示单位矩阵。

由公式(2)可知此分类器只能解决线性输入问题, 为解决非线性输入问题, 利用核函数  $w = \sum_i \alpha_i x_i$  将其转换到对偶空间下, 并利用快速傅里叶变换和循环矩阵性质将时域内计算转换到频域, 得到:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{x}^* \odot \hat{x} + \lambda}, \quad (3)$$

式中: 上标表示傅里叶变换, 上标 \* 表示共轭矩阵, 符号  $\odot$  表示矩阵对应位置元素的点乘,  $\alpha$  表示对偶空间下的分类器参数。

在跟踪过程中利用公式(3)得到分类器参数, 然后进行目标快速响应检测, 其定义为:

$$f(z) = F^{-1}(\hat{z} \odot \hat{x}^* \odot \hat{\alpha}), \quad (4)$$

式中:  $F^{-1}$  表示傅里叶逆变换,  $z$  表示当前跟踪帧的输入图像,  $f(z)$  为计算得到的时域内输入图像响应图, 在响应图中最大响应所在位置就是跟踪目标所在位置。

### 3 本文算法

为了解决嘈杂背景下的目标跟踪问题, 本文提出基于视觉显著性和扰动模型的上下文感知跟踪, 下面为本文算法的详细介绍。

#### 3.1 上文感知相关跟踪

在相关滤波跟踪中跟踪目标容易受到背景干扰影响, 目标周围的上下文信息对提高跟踪算法鲁棒性有重要作用, 在上下文感知跟踪<sup>[12]</sup>中, 将目标周围的上下文信息引入到相关滤波分类器训练过程中, 构造上下文感知相关跟踪器。在上下文感知相关跟踪算法中, 在目标周围采样  $n$  个上

$$\hat{\alpha} = \begin{bmatrix} \text{diag}(\hat{x} \odot \hat{x}^*) + \lambda_1 & \cdots & \text{diag}(\sqrt{\lambda_2} \hat{x} \odot \hat{x}^n)^* \\ \vdots & \ddots & \vdots \\ \text{diag}(\sqrt{\lambda_2} \hat{x}^n \odot \hat{x}^*) & \cdots & \text{diag}(\lambda_2 \hat{x}^n \odot (\hat{x}^n)^*) + \lambda \end{bmatrix}^{-1} \begin{bmatrix} \hat{y} \\ \vdots \\ 0 \end{bmatrix} \quad (11)$$

由公式(11)可知所有子数据块均可对角化, 说明任意两个子数据块是相互独立的, 因此可以将此系统分解为数个子系统进行求解, 此时计算得到分类器参数  $\alpha = [\alpha_0, \dots, \alpha_k]$ , 将数据参数代入

下文图像块  $\hat{x}^j$ , 目标搜索窗口区域为  $x$ , 图像块  $\hat{x}^j$  和  $x$  所对应的循环数据矩阵为  $\tilde{\mathbf{X}}^j$  和  $\mathbf{X}$ , 则公式(1)变为:

$$\epsilon = \underset{\omega}{\text{argmin}} \sum_i \|x_i \omega - y_i\|_2^2 + \lambda_1 \|\omega\|_2^2 + \lambda_2 \sum_j \sum_i \|\hat{x}_i^j \omega\|_2^2, \quad (5)$$

式中  $\lambda_1$  和  $\lambda_2$  均表示正则化因子。

为了推导方便, 将公式(5)改写为:

$$\epsilon = \underset{\omega}{\text{argmin}} \sum_i \|\mathbf{B}_i \omega - \mathbf{y}'\|_2^2 + \lambda_1 \|\omega\|_2^2, \quad (6)$$

式中:  $\mathbf{B}_i = \begin{bmatrix} x_i \\ \sqrt{\lambda_2} \hat{x}_i^j \\ \vdots \\ \sqrt{\lambda_2} \hat{x}_i^n \end{bmatrix}, \mathbf{y}' = \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \hat{x}_i^j$  表示上下文图

像样本的  $\hat{x}^j$  的循环偏移样本。

对公式(6)求闭环求解, 得到分类器参数为:

$$\omega = \frac{\mathbf{B}^T \mathbf{y}'}{\mathbf{B}^T \mathbf{B} + \lambda_1 \mathbf{I}}, \quad (7)$$

式中  $\mathbf{B}$  表示由所有循环偏移样本  $\mathbf{B}_i$  组成的循环数据矩阵。

将循环数据矩阵  $\mathbf{B}$  和  $\mathbf{B}^T$  代入公式(7)中, 并由循环矩阵的性质  $\mathbf{X} = F \text{diag}(\hat{x}) F^H$  和  $\mathbf{X}^T = F \text{diag}(\hat{x}^*) F^H$  对公式(7)进行化简:

$$\hat{\omega} = \frac{\hat{x}^* \odot \hat{y}'}{\hat{x}^* \odot \hat{x} + \lambda_1 + \lambda_2 \sum_{j=1}^n \hat{x}^j \odot \hat{x}^j}. \quad (8)$$

此时利用分类器  $\hat{\omega}$  实现快速目标检测, 得到目标响应为:

$$f(z)^{\text{cf}} = F^{-1}(\hat{z} \odot \hat{\omega}). \quad (9)$$

同理类似于公式(3)可以得到对偶空间下分类器定义:

$$\alpha = \frac{\mathbf{y}'}{\mathbf{B} \mathbf{B}^T + \lambda_1 \mathbf{I}} \quad (10)$$

利用循环矩阵的性质, 将矩阵  $\mathbf{B}$  和  $\mathbf{B}^T$  代入公式(10)化简得到:

$$\hat{\alpha} = \begin{bmatrix} \text{diag}(\sqrt{\lambda_2} \hat{x} \odot \hat{x}^n)^* \\ \vdots \\ \text{diag}(\lambda_2 \hat{x}^n \odot (\hat{x}^n)^*) + \lambda \end{bmatrix}^{-1} \begin{bmatrix} \hat{y} \\ \vdots \\ 0 \end{bmatrix} \quad (11)$$

$f(z) = \mathbf{Z} \mathbf{B}^T \alpha$  中化简得到:

$$\hat{f}(z)^{\text{cf}} = \hat{z} \odot \hat{x}^* \odot \hat{\alpha}_0 + \sqrt{\lambda_2} \sum_{j=1}^k \hat{z} \odot \hat{x}^j \odot \hat{\alpha}_j. \quad (12)$$

由公式(12)求得对偶空间下频域内输入图像

的响应,利用快速傅里叶变换将其转换到时域内,最大响应所在位置就是跟踪目标所在位置。

### 3.2 直方图扰动模型

为了提高形变情况下跟踪算法的鲁棒性,引入直方图扰动模型。相关滤波算法对目标形变较为敏感,相关滤波跟踪通常利用局部特征构建目标外观模型(如,HOG 特征、灰度特征),因此对目标外观模型的局部变化较为敏感,而直方图模型利用的是统计特征,虽然直方图跟踪对光照变化较为敏感,但是其对目标形变具有较好的跟踪鲁棒性,利用直方图扰动模型能有效提高复杂背景下的相似目标干扰问题,下面对直方图扰动模型进行介绍:

首先建立目标图像的颜色概率模型,利用贝叶斯分类器对目标像素和背景像素进行建模,得到:

$$P(c \in O | O, B, b_c) = \frac{P(b_c | c \in O) P(c \in O)}{\sum_{\Omega} P(b_c | c \in \Omega) P(c \in \Omega)}, \quad (13)$$

式中: $O$ 和 $B$ 分别表示目标像素区域与背景像素区域, $P(b_c | c \in O)$ 表示 $O$ 中像素 $c$ 属于区间 $b_c$ 的直方图概率, $P(b_c | c \in \Omega)$ 表示像素空间 $\Omega$ 中像素属于区间 $b_c$ 的概率, $P(c \in O | O, B, b_c)$ 为构造的贝叶斯分类器模型,其表示像素 $c$ 属于目标像素的概率。

为了更好地理解公式(13),对公式(13)进行重写得到:

$$P(c \in O | O, B, b_c) = \begin{cases} \frac{H_o^1(b_c)}{H_o^1(b_c) + H_b^1(b_c)} & \text{if } (x(c) \in (O, B)) \\ 0 & \text{otherwise} \end{cases}, \quad (14)$$

式中: $H_o^1(\cdot)$ 和 $H_b^1(\cdot)$ 分别表示目标与背景的颜色直方图,且满足 $P(b_c | c \in O) \approx H_o^1(b_c) / |O|$ , $P(b_c | c \in B) \approx H_b^1(b_c) / |B|$ ,这里 $|O|$ 和 $|B|$ 表示目标区域与背景区域内的像素个数。

公式(14)是关于目标与背景的颜色概率模型,为了提高算法的抗扰动能力,再引入目标-扰动模型,其定义如下:

$$P(c \in O | O, D, b_c) = \begin{cases} \frac{H_o^1(b_c)}{H_o^1(b_c) + H_b^1(b_c)} & \text{if } (x(c) \in (O, D)) \\ 0 & \text{otherwise} \end{cases}, \quad (15)$$

式中: $H_b^1(\cdot)$ 表示扰动区域的颜色直方图模型。

通过公式(14)和(15)得到最终的目标联合概率模型:

$$P(c \in O | b_c) = \eta \times P(c \in O | O, B, b_c) + (1 - \eta) P(c \in O | O, D, b_c). \quad (16)$$

最终由目标联合概率直方图得到的响应图定义为:

$$f(z)^{\text{hist}} = \frac{1}{|W_z|} \sum_{c_{i,j} \in W_z} P_t(c_{i,j}), \quad (17)$$

式中: $|W_z|$ 表示搜索窗口 $z$ 内的像素个数, $P_t(c_{i,j}) = P(c_{i,j} \in O | b_{c_{i,j}})$ ,表示第 $t$ 帧时像素 $c_{i,j}$ 属于目标像素的概率值, $f(z)^{\text{hist}}$ 为表示输入图像 $z$ 的颜色直方图概率响应。

### 3.3 基于图像签名的视觉显著性图

为了解决遮挡情况下的目标跟踪问题,本文利用显著性检测构建目标重检测模块。本文利用图像签名算法<sup>[13]</sup>获取输入图像的稀疏显著性图,图像签名算法定义为:

$$\text{ImageSignature}(x) = \text{sign}[\text{DCT}(x)], \quad (18)$$

式中: $\text{DCT}(\cdot)$ 表示离散余弦变换, $\text{sign}(\cdot)$ 表示符号函数, $\text{ImageSignature}(\cdot)$ 表示输入图像的签名图像。

文献[13]说明利用图像签名算法重构图像可以获得图像的显著性图(如图1所示),图像签名重构图像定义如式(19)。



(a) 输入图像 (b) 图像签名重构图  
(a) Input image (b) Reconstruction image of image signature



(c) 显著性图 (d) 显著性图可视化  
(c) Saliency map (d) Saliency map visualization

图1 稀疏显著性检测

Fig. 1 Sparse salient detection

$$\hat{x} = IDCT(\hat{x}), \quad (19)$$

式中:  $\hat{x} = DCT(x)$ ,  $\hat{x}$  即为利用图像签名算法重构的图像。

利用图像签名重构图像计算稀疏显著性图, 有:

$$x^{\text{salient}} = g * (\hat{x} \odot \hat{x}), \quad (20)$$

式中:  $x^{\text{salient}}$  为目标图像稀疏显著性图,  $g$  为高斯平滑函数,  $*$  表示矩阵的卷积操作。

### 3.4 目标位置估计与模板更新

为了提高目标跟踪鲁棒性, 将上下文相关跟踪响应与直方图扰动模型响应进行加权融合, 得到最终的目标响应图:

$$f(z)' = \delta f^{\text{cf}}(z) + (1 - \delta) f^{\text{hist}}(z), \quad (21)$$

式中:  $\delta$  表示响应权重因子,  $f(z)'$  是目标的联合响应图, 其最大响应所在位置就是跟踪目标所在位置。

在跟踪过程中利用峰旁比<sup>[7]</sup>对目标响应置信度进行估计, 以此判别跟踪目标是否发生遮挡, 当目标置信度低于指定阈值, 利用稀疏显著性图获得候选目标  $V = \{v_1, \dots, v_k\}$ , 然后计算不同候选目标的相关滤波响应, 响应最大的候选目标即为跟踪目标所在位置, 从而解决目标重定位问题。这里判断目标是否发生遮挡利用如下公式:

$$\phi = \begin{cases} 1 & R_t^{\text{psr}} > R_{\text{threshold}}^{\text{psr}} \\ 0 & \text{otherwise} \end{cases}, \quad (22)$$

式中:  $R_t^{\text{psr}}$  表示当前帧目标响应峰旁比,  $R_{\text{threshold}}^{\text{psr}}$  表示峰旁比阈值,  $\phi$  表示目标是否发生遮挡, 当为 1 时表示跟踪正常, 为 0 时表示出现跟踪遮挡。

采用自适应更新策略对目标模板和直方图模型进行更新:

$$x_t = (1 - \phi\varphi) x_{t-1} + \phi\varphi \hat{x}_t, \quad (23)$$

$$P(c \in O|b_c)_t = (1 - \phi\varphi) P(c \in O|b_c)_{t-1} + \phi\varphi \dot{P}(c \in O|b_c)_t, \quad (24)$$

式中:  $\varphi$  表示模板学习速率,  $x_t$  和  $P(c \in O|b_c)_t$  表示第  $t$  帧学习得到的目标模板和直方图模型,  $\hat{x}_t$  和  $\dot{P}(c \in O|b_c)_t$  表示第  $t$  帧采样得到的目标样本和直方图模型。

同时对公式(7)中的相关滤波分类器分子和分母进行更新:

$$\mathbf{B}_t^T y' = (1 - \phi\varphi) \mathbf{B}_{t-1}^T y' + \phi\varphi \dot{\mathbf{B}}_t y', \quad (25)$$

$$\mathbf{B}_t^T \mathbf{B}_t + \lambda = (1 - \phi\varphi) (\mathbf{B}_{t-1}^T \mathbf{B}_{t-1} + \lambda) + \phi\varphi (\dot{\mathbf{B}}_t^T \dot{\mathbf{B}}_t + \lambda), \quad (26)$$

式中:  $\mathbf{B}_t$  表示第  $t$  帧学习得到的循环样本数据矩

阵,  $\dot{\mathbf{B}}_t$  表示第  $t$  帧采样得到的循环样本数据矩阵。

同时为了解决目标尺度变化问题, 本文采用尺度池策略<sup>[14]</sup>对目标图像进行不同尺度采样, 获取不同尺度大小的目标图像  $z_{s_i}$ ,  $s_i$  表示尺度因子, 然后对不同尺度图像进行线性插值, 利用相关滤波算法计算不同尺度目标图像块响应, 最大响应所在尺度就是当前目标尺度变化情况。

## 4 实验结果与分析

### 4.1 实验环境与参数设置

本文所有实验均是在 MATLAB2014b 软件上进行测试, 实验电脑配置为 Intel i5 3.2 GHz CPU, 16 G 内存。实验参数如下: 相关滤波算法搜索窗口为目标大小的 2.0 倍大, 目标联合概率模型中权重因子  $\eta$  设置为 0.5, 加权响应权重因子  $\delta$  设置为 0.5, 目标模型学习速率  $\gamma$  设置为 0.025, 学习速率  $\varphi$  设置为 0.01, 尺度池大小为 7, 候选目标个数设置为 10, 正则化因子  $\lambda$  设置为  $10^{-4}$ 。为了更好地分析本文算法性能的优劣性, 利用 OTB 公开数据集测试本文算法并与其它 8 种算法进行比较。

### 4.2 实验结果分析

为了验证本文算法有效性, 利用数据集<sup>[11]</sup>中的视频序列进行测试(测试视频属性如表 1 所示), 主要测试算法在背景嘈杂 (Background Clutter, BC)、遮挡 (Occlusion, OCC)、形变 (Deformation, DEF)、快速运动 (Fast Motion, FM) 和尺度变化 (Scale Variation, SV) 等情况下性能表现。同时与 LCT、MTT、DSST、KCF、TLD<sup>[15]</sup>、STC、TGPR<sup>[16]</sup> 和 SCM<sup>[17]</sup> 等 8 种流行算法进行对比分析。

表 1 测试视频属性

Tab. 1 Properties of experimental video

测试视频	帧数	图像大小	视频属性
boy	602	640 × 480	FM, SV
box	1161	640 × 480	OCC, BC, SV
bolt	350	640 × 360	DEF, OCC
jogging-2	307	352 × 288	OCC, DEF
doll	3872	400 × 300	SV, OCC
bolt2	293	480 × 270	DEF, BC
blurCar2	585	640 × 480	SV, FM
blurOwl	631	640 × 480	FM, SV
board	698	640 × 480	BC, SV, FM
basketball	725	576 × 432	BC, OCC, DEF

为了更好分析本文算法性能,利用中心位置误差(Center Location Error, CLE)、距离精确度(Distance Precision, DP)、重叠率精度(Overlap Precision, OP)、精确度图和成功率图等<sup>[11]</sup>评价标准测试算法性能。

4.2.1 整体性能分析

利用表 1 中的视频序列测试所提算法在复杂场景下的整体跟踪性能,图 2 为 9 种算法在测试视频序列上的整体精确度和成功率图。由图 2 可知本文算法的精确度得分和成功率得分在 9 种算

法中均排名第一,本文算法的成功率得分为 0.695,相比排名第二的 LCT(0.615)跟踪算法提升 13.0%,本文算法的精确度得分为 0.708,相比排名第二的 DSST(0.571)跟踪算法提升 24.0%。通过实验说明本文算能较好的解决尺度变化、快速运动、形变、背景嘈杂和遮挡等复杂情况下的跟踪问题,相比传统的相关滤波算法在成功率和精确度上均有较大幅度的提升。另外,经过测试本文算法在测试序列上平均跟踪速度为 22.8 f/s,基本满足算法实时性要求。

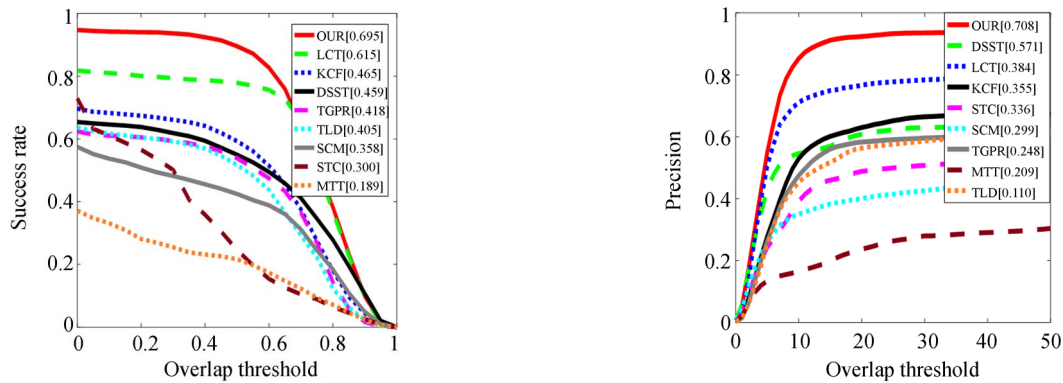


图 2 9 种算法的成功率图和精确度图

Fig. 2 Success plots and precision plots of nine algorithm

表 2 9 种算法在测试序列上的距离精确度

Tab. 2 Distance precision of nine algorithms on test sequences

Video	OUR	LCT	KCF	DSST	TGPR	MTT	STC	TLD	SCM
boy	1.000	1.000	1.000	1.000	1.000	0.706	0.761	1.000	0.440
box	0.424	0.068	0.415	0.394	0.294	0.259	0.390	0.742	0.153
bolt	1.000	1.000	0.989	1.000	0.026	0.017	0.046	0.306	0.031
jogging-2	1.000	0.974	0.163	0.186	0.997	0.173	0.186	0.857	1.000
doll	0.992	0.981	0.968	0.993	0.971	0.610	0.763	0.983	0.978
bolt2	0.991	0.017	0.017	0.020	0.020	0.017	0.980	0.014	0.014
blurCar2	1.000	0.998	0.938	1.000	0.962	0.133	0.990	0.979	0.099
blurOwl	0.968	0.891	0.228	0.233	0.512	0.114	0.114	0.713	0.190
board	0.902	0.734	0.656	0.791	0.053	0.049	0.092	0.011	0.441
basketball	0.973	1.000	0.923	0.462	0.993	0.288	0.560	0.028	0.661
All	0.708	0.384	0.355	0.571	0.248	0.2090	0.336	0.110	0.299

表 3 9 种算法在测试序列上的重叠率精度

Tab. 3 Overlap precision of nine algorithms on test sequences

Video	OUR	LCT	KCF	DSST	TGPR	MTT	STC	TLD	SCM
boy	1.000	1.000	0.992	0.993	0.990	0.487	0.663	0.935	0.439
box	0.432	0.090	0.358	0.337	0.241	0.264	0.128	0.760	0.176
bolt	0.998	0.989	0.934	1.000	0.014	0.011	0.043	0.146	0.014
jogging-2	0.998	0.971	0.160	0.182	0.990	0.163	0.173	0.931	0.990
doll	0.998	0.994	0.552	0.730	0.714	0.504	0.101	0.624	0.987
bolt2	0.772	0.007	0.007	0.010	0.010	0.010	0.515	0.007	0.007
blurCar2	0.953	1.000	0.947	1.000	0.932	0.133	0.311	1.000	0.144
blurOwl	0.958	0.894	0.228	0.235	0.512	0.119	0.074	0.735	0.216
board	0.899	0.852	0.854	0.861	0.129	0.191	0.219	0.107	0.663
basketball	0.958	0.992	0.898	0.148	0.850	0.289	0.236	0.025	0.611
ALL	0.895	0.779	0.593	0.550	0.538	0.217	0.246	0.517	0.425

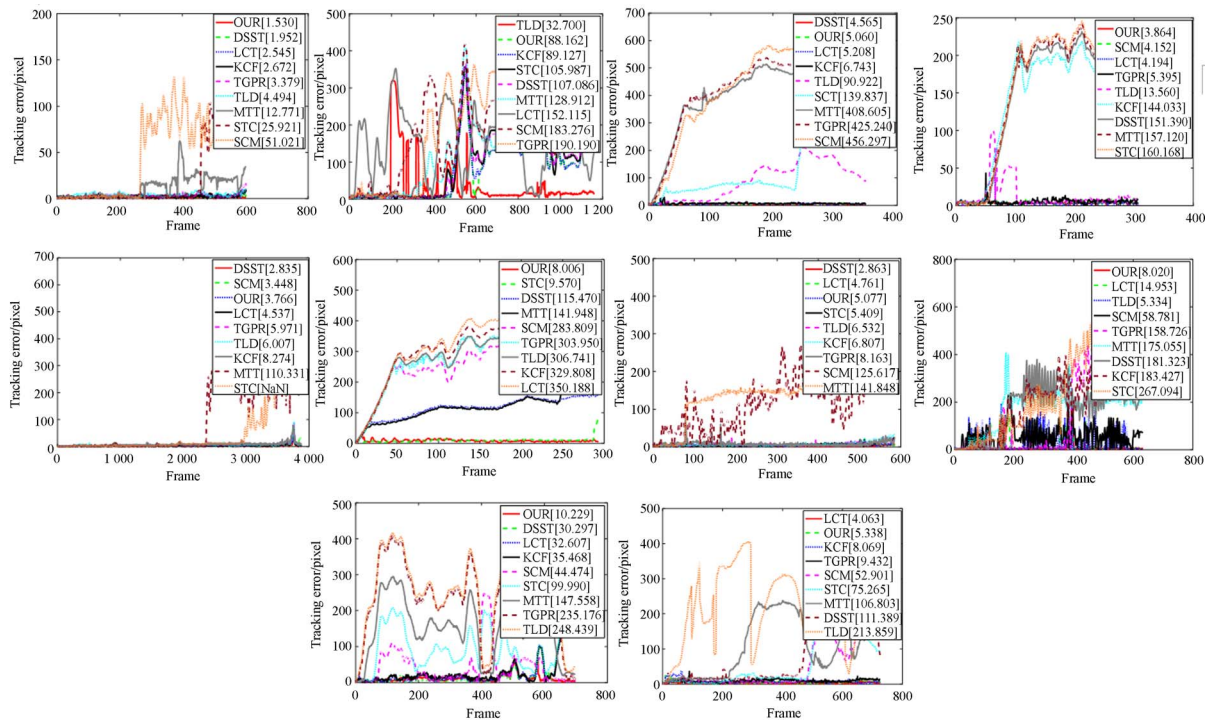


图 3 9 种算法在测试序列上的跟踪误差曲线

Fig. 3 Tracking error curves of nine algorithms on test sequences

同时记录算法在不同测试序列上的距离精确度和重叠率精度(如表 2 与表 3 所示)以及跟踪误差曲线(如图 3 所示)。从表 2 与表 3 中看出,本文算法的整体距离精确度和重叠率精确度分别为 0.708(中心位置误差阈值设置为 20 个像素)和 0.895(重叠率阈值设为 0.5),在 9 种算法中均排名第一,且本文算法在 10 个测试序列中的 7 个序列上跟踪距离精确度排名第一,在 5 个序列上重叠率精确度排名第一。另由图 3 可知,本文算法

在 10 个视频序列中的 5 个序列上平均跟踪误差最小,且在所有序列上平均跟踪误差均排名前三,这说明本文算法具有较高的跟踪精度,能有效提高复杂场景下跟踪器的鲁棒性。

#### 4.2.2 基于视频属性分析

为了更好地分析本文算法,分别测试并记录算法在背景嘈杂、遮挡、形变、快速运动和尺度等情况下的实际跟踪情况(如图 4 和图 5 所示),下面算法在不同场景下性能分析。

### 实验一:背景嘈杂情况下性能分析

在 basketball 和 board 视频序列中,跟踪目标处于嘈杂的环境中,此时跟踪器受嘈杂背景的影响,容易发生跟踪漂移或者跟踪失败。例如在 board 序列的 #40、#267 和 #509 帧中,受嘈杂背景的影响,STC、TLD、TGPR 和 MTT 算法均发生不同程度的跟踪失败现象,而本文算法仍能正确跟踪,这是因为本文算法利用直方图扰动模型和上下文感知相关滤波器显著提高了嘈杂背景下分类器分类能力,且从图 4 中得出本文算法在 board 序列上平均跟踪误差为 10.229,优于 KCF 算法的 35.468。在 basketball 序列上,本文算法的平均跟踪误差也小于 KCF 算法,而 LCT 算法因为利用了多种特征构建目标外观模型,因此跟踪误差也较小。

### 实验二:遮挡情况下性能分析

在 jogging-2 和 box 序列上,跟踪目标被遮挡物部分遮挡或者完全遮挡,此时容易造成跟踪失败。在跟踪目标发生遮挡的情况下,首先目标外

观模型将发生误更新,其造成后续跟踪漂移问题,同时如果算法不具备重检测功能,无法解决完全遮挡情况下目标重定位问题。例如在 jogging-2 序列中,由 #39 到 #80 帧,行人被电线杆完全遮挡,此时相关滤波类算法中只有本文算法与 LCT 算法跟踪成功,而 KCF 与 DSST 算法因为不具备重检功能均跟踪失败。

### 实验三:形变情况下性能分析

在 bolt2 和 bolt 序列中,运动员均发生剧烈形变,此时目标外观模型发生剧烈变化,跟踪器容易分类错误,从而跟踪失败。例如在 bolt2 序列 #13、#49 和 #147 帧中,随着目标不断发生形变,最终只有本文算法和 STC 算法能一直跟踪指定目标,STC 算法因为利用了目标周围的上下文信息,因此减缓目标形变对跟踪算法的影响。相关滤波算法 KCF 因为只利用局部特征构建目标外观模型,因此其对形变较为敏感,而本文算法将相关滤波算法与贝叶斯分类进行有效结合,利用直方图统计信息构建目标模型,降低了形变对算法影响。

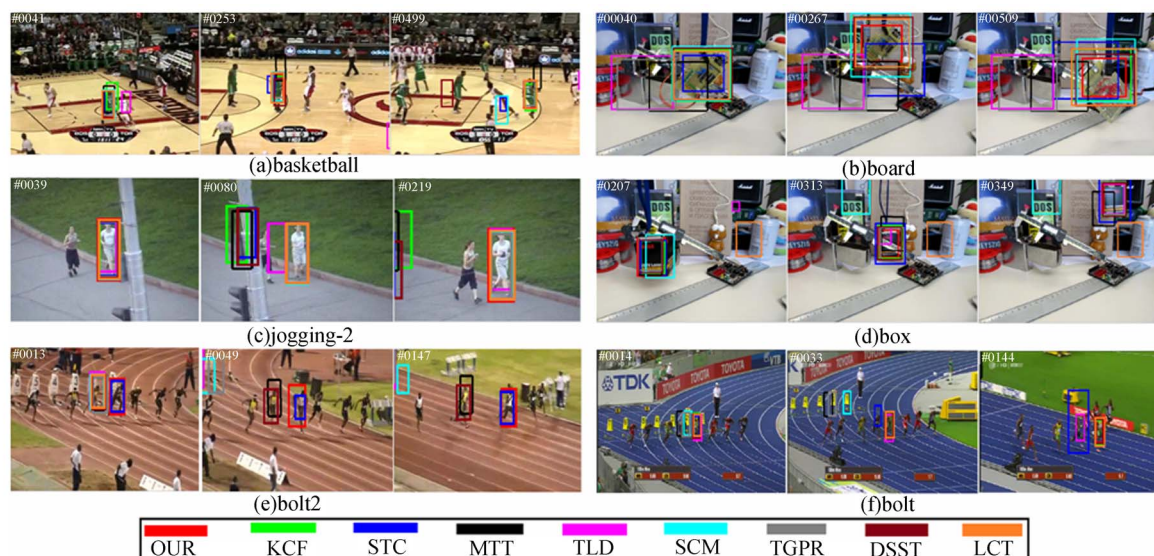


图 4 背景嘈杂、遮挡和形变情况下的跟踪结果

Fig. 4 Tracking results under background noise, occlusion and deformation

### 实验四:快速运动情况下性能分析

在 blurOwl 和 boy 中跟踪目标发生快速运动,受跟踪算法搜索空间的限制,在快速运动情况下容易产生跟踪失败。例如在 blurOwl 序列中,受相机抖动的影响,跟踪目标发生剧烈的运动,此时只有 LCT、TLD 和本文算法能正确跟踪,因为

LCT 算法与 TLD 算法均具有长期跟踪模块,因此能解决快速运动问题,而本文算法因为利用视觉显著性检测构建了目标重检测模块,因此本算法也能解决目标快速运动问题。

### 实验五:尺度变化情况下性能分析

在 doll 和 bulrCar2 序列上,目标发生不同程

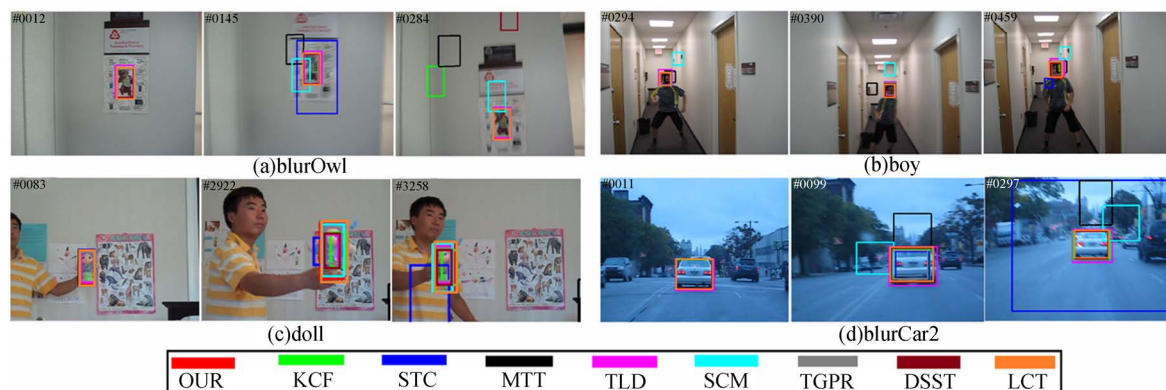


图 5 快速运动和尺度变化情况下的跟踪结果

Fig. 5 Tracking results under fast motion and scale variation

度尺度变化。例如在 doll 序列的 # 83、# 2922 和 # 3258 帧中,玩具熊不断发生尺度变化,而 KCF 算法因为不具有尺度估计策略,因此无法解决尺度估计问题,STC 算法虽然具有尺度估计模块,但是其发生了剧烈的尺度漂移问题,而本文算法采用尺度估计策略,因此能很好解决尺度估计问题。

#### 4.2.3 算法贡献度分析

本文算法的基础是相关滤波算法,为了更好的分析本文算法相比传统相关滤波算法的增益度,分别记录本文算法、KCF+CA(CA 即表示上下文感知模型)和 KCF 的成功率得分(如表 4 所示)。从表 4 中得出,本文算法利用 HOG 特征和直方图特征构建目标外观模型,且支持目标尺度变化,本文算法的成功率得分为 0.695,相比 KCF 算法和 KCF+CA 算法分别提升了 28.2% 和 49.4%,这说明本文引入视觉显著性检测和直方图扰动模型能有效解决遮挡、背景嘈杂和形变等

表 4 算法属性与成功率

Tab. 4 Algorithm attributes and success ratio

跟踪算法	特征	尺度	成功率
OUR	Color histogram, HOG	Yes	0.695
KCF+CA	HOG	No	0.542
KCF	HOG	No	0.465

#### 参考文献:

[1] 王晔,王春平,李军,等. 特征融合和模型自适应更

复杂情况下的跟踪失败问题,显著提高算法跟踪鲁棒性。

## 5 结 论

针对相关滤波跟踪算法中存在的背景嘈杂、尺度变化和遮挡问题,提出利用视觉显著性和扰动模型的上下文感知相关跟踪。为了解决遮挡情况下的目标重检测问题,利用图像签名算法构建了目标重检测模块,提高了算法的跟踪鲁棒性;同时,将目标周围的背景信息引入到分类器训练中,构建了上下文相关滤波跟踪,提高了嘈杂背景下分类器的分类能力;然后,将相关滤波响应与直方图滤波响应进行加权融合,以此解决目标形变情况下的鲁棒目标跟踪问题;最后,利用尺度池策略和自适应模板更新策略解决目标尺度估计问题和遮挡情况下的模板漂移问题。实验表明本文算法在精确度和成功率上得分分别为 0.695 和 0.708,均优于传统相关滤波算法,能很好的解决复杂背景下的精确目标跟踪问题。为了进一步解决目标旋转问题并提高算法的跟踪精度,下一阶段研究重点是基于深度卷积神经网络的研究,利用深度卷积神经网络可以解决目标旋转问题并大幅提高算法精确度。

新相结合的相关滤波目标跟踪[J]. 光学 精密工程, 2016, 24(8): 2059-2066.

WANG W, WANG CH P, LI J, *et al.*. Correlation filter tracking based on feature fusing and mod-

- el adaptive updating [J]. *Opt. Precision Eng.*, 2016, 24(8):2059-2066. (in Chinese)
- [2] 王春平,王晔,刘江义,等. 基于色度饱和度-角度梯度直方图特征的尺度自适应核相关滤波跟踪[J]. *光学精密工程*, 2016, 24(9):2293-2301.  
WANG CH P, WANG W, LIU J Y, *et al.*. Scale adaptive kernelized correlation filter tracking based on HHS-HOG feature [J]. *Opt. Precision Eng.*, 2016, 24 (9):2293-2301. (in Chinese)
- [3] 张红颖,李灿锋. 结合特征在线选择与协方差矩阵的压缩跟踪算法[J]. *光学精密工程*, 2017, 25(4):519-527.  
ZHANG H Y, LI C F. Compressive tracking algorithm combining online feature selection with covariance matrix [J]. *Opt. Precision Eng.*, 2017, 25 (04):519-527. (in Chinese)
- [4] POSSEGER H, MAUTHNER T, BISCHOF H. In defense of color-based model-free tracking [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2015: 2113 -2120.
- [5] 吉训生,陈赛,黄越. 判别稀疏表示与在线字典学习的运动目标跟踪[J]. *计算机工程与应用*, 2017, 53(3):211-215.  
JI X SH, CHEN S, HUANG Y. Discriminative sparse representation and online dictionary learning for visual tracking [J]. *Computer Engineering and Applications*, 2017, 53(3):211-215. (in Chinese)
- [6] ZHANG T, GHANEM B, LLU S, *et al.*. Robust visual tracking via structured Multi-Task sparse learning [J]. *International Journal of Computer Vision*, 2013, 101(2):367-383.
- [7] BOLME D S, BEVERIDGE J R, DRAPER B A, *et al.*. Visual object tracking using adaptive correlation filters [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2010: 2544-2550.
- [8] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.*. High-speed tracking with kernelized correlation filters [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 583-596.
- [9] ZHANG K, ZHANG L, LIU Q, *et al.*. Fast visual tracking via dense spatio-temporal Context Learning [C]. *European Conference on Computer Vision*, Springer, Cham, 2014: 127-141.
- [10] MA CH, YANG X K, ZHANG CH Y, *et al.*. Long-term correlation tracking [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2015: 5388-5396.
- [11] WU Y, LIM J W, YANG M H. Online object tracking: a benchmark [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2013: 2411-2418.
- [12] MUELLER M, SMITH N, GHANEM B. Context-aware correlation filter tracking [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2017:1-9.
- [13] HOU X, HAREL J, KOCH C. Image signature: highlighting sparse salient regions [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(1):194-201.
- [14] DANELLJAN M, H? GER G, KHAN F S, *et al.*. Accurate scale estimation for robust visual tracking [C]. *British Machine Vision Conference*, IEEE, 2014:1-5.
- [15] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking -learning-detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7):1409-1422.
- [16] GAO J, LING H, HU W, *et al.*. Transfer learning based visual tracking with gaussian processes regression [C]. *European Conference on Computer Vision*, Springer, Cham, 2014: 188-203.
- [17] ZHONG W, LU H, YANG M H. Robust object tracking via sparsity-based collaborative model [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2012:1838-1845.

#### 作者简介:



张 博(1980—),男,湖南长沙人,硕士,高级实验师,2009年于湖南师范大学获得硕士学位,现为长沙师范学院信息与工程实验中心主任,主要从事智能感知与控制、目标检测与跟踪方面的研究。E-mail: zb801121@ 126. com