

一种新的基于颜色统计特征的判别跟踪方法^①

胡昭华^{②*} 鞠 蓉^{③*} 欧阳雯* 金 蓉*

(* 南京信息工程大学 电子与信息工程学院 南京 210044)

(** 南京信息工程大学 江苏省大气环境与装备技术协同创新中心 南京 210044)

摘 要 针对目前目标跟踪的大部分判别算法注重跟踪效率而没有考虑尺度变化这一问题,提出了一个简单而鲁棒的基于颜色统计特征的判别跟踪方法。这种新的颜色统计特征具有一定的光照不变性,同时保持较强的判别能力。建立了跟踪过程的仿射运动模型,利用优化参数来解决尺寸及角度变化等问题。此外,为了进一步提高跟踪速度,采用低维的颜色统计特征描述目标外观,利用颜色统计特征训练贝叶斯分类器,将置信值最大的样本作为跟踪结果,并在线更新分类器。与现有跟踪器的大量综合性的对比实验表明,该判别跟踪方法在不同挑战因素下均有明显优势。

关键词 视觉跟踪, 颜色统计特征, 仿射变换, 降维, 判别分类器

0 引 言

目标跟踪在计算机视觉的许多应用中都占据非常重要的地位,如自动监测、视频检索、交通监控和人机交互等。尽管研究者们在过去几十年里提出了很多目标跟踪算法,但要构建一个简单且鲁棒的跟踪系统用来处理因突然运动、光照变化、形状变形和遮挡等引起的外观变化,依然具有很大的挑战。本研究针对这种情况进行了探索,提出了一种新的基于颜色统计特征的判别跟踪算法,实验证明,该算法的性能明显优于其他算法。

1 相关研究

选择合适的视觉特征对于一个高性能的跟踪器是很重要的。大部分先进的跟踪器依赖于灰度、形状、强度和纹理信息^[1-5]。方向梯度直方图(HOG)特征^[6]描述了目标的边缘特征。局部二值模式

(LBP)特征^[7]不容易受到光照的影响,可以提取图像的局部纹理特征,但它对噪声敏感,容易受到目标变形的影响。基于 Haar-like 特征的跟踪器^[8-10]在跟踪速度和检测率上占有优势,然而对边缘却很敏感。现在越来越多的研究者在研究颜色特征。与其他的视觉特征相比,颜色特征具有较高的鲁棒性,对目标大小、方向和视角的依赖性更小。Swain 和 Ballard^[11]提出用颜色直方图(color histogram)提取图像的特征。虽然图像直方图受图像的物理变换影响较小,然而,由于颜色直方图描述的是全局颜色统计的结果,丢失了像素之间的位置信息,因而在识别前景对象上不能取得显著的效果。为了消除颜色直方图中量化的影响,Stricker 和 Orengo^[12]提出用颜色矩(color moments)作为图像检索中的特征向量。然而,为避免低次矩较弱的分辨能力,颜色矩常和其它特征结合使用。颜色直方图和颜色矩只包含每个像素的颜色信息,忽略了局部图像相邻像素之间的关系。Huang 等人^[13]提出了颜色相关图(color cor-

① 国家自然科学基金青年基金(61203273),江苏省自然科学基金青年基金(BK20141004),江苏省普通高校自然科学研究资助项目(11KJB510009, 14KJB510019),江苏省“信息与通信工程”优势学科资助和江苏省大学生实践创新训练计划(201410300021)资助项目。
② 女,1981年生,博士,副教授;研究方向:视频目标跟踪,模式识别,粒子滤波等;E-mail: zhaohua_hu@163.com
③ 通讯作者,E-mail: jrmuist@163.com
(收稿日期:2015-12-01)

relogram),既直观地描述了不同颜色的像素数量在整幅图像中所占的比例,还反映了不同颜色对间的空间相关性。Joost 等人^[14]从现实图像中提出了一种新的色名(color name)特征。由于光源、阴影、高光、相机和目标形状的变化,一个图像序列的颜色测量会有很大的变化,通过对彩色成像中这些因素的鲁棒性研究,色名特征已成功应用于图像分类和行为识别中。虽然颜色信息在视觉跟踪上得到了重大进展,但仅限在简单的颜色空间转换上。相比于视觉跟踪,复杂的颜色特征在目标识别和检测方面表现出更好的性能。将颜色信息用于视觉跟踪是一个困难的挑战。Martin 等人^[15]将颜色属性(color attributes)特征引入到视觉跟踪中,展现了该特征优异的光照不变性和判别能力。受上述工作的启发,本文提出了一种新的基于颜色特征的跟踪方法。

跟踪算法一般可以分成两类:生成模型^[2,3]和判别模型^[4,5,8-10]。生成算法通过学习一个外观模型来搜索图像,将重构误差最小的区域作为跟踪结果。这类算法都是基于模板或子空间模型。Frag 跟踪器^[1]采用局部图像块的直方图,有效地解决了部分遮挡问题。然而,由于没有有效的模板更新策略,很难处理目标的外观和尺寸变化。视频跟踪分解(visual tracking decomposition, VTD)跟踪器^[2]则是利用多个运动和观测模型来描述目标外观和尺寸的变化,从而有效扩展了传统的粒子滤波框架。增量视觉跟踪器(incremental visual tracker, IVT)^[3]能够适应光照和姿态变化造成的外观变化,这是因为它

在线增量学习低维子空间。然而,当目标经历严重的遮挡时,跟踪器的性能却不尽如人意。生成模型没有考虑到目标周围环境的有用信息,这在一定程度上限制了生成跟踪算法的鲁棒性。

判别方法的目标是将目标从背景中区分出来,它将目标跟踪转换成一个通过训练分类器区分目标和其周围背景的检测问题。Avidan 等人^[16]利用 Adaboost 对每个像素进行分类,训练一系列弱分类器级联成一个强分类器来区分目标和背景。该方法可以有效地减少遮挡造成的影响。由于自主学习导致的积累不准确性,Babenko 等人^[9]提出用多示例学习(multiple instance learning, MIL)代替传统的监督学习。虽然这种方法可以减少跟踪漂移,但不能处理非刚性变形。判别跟踪算法很简单,而且实验结果具有良好的鲁棒性,但大多数跟踪器无法从根本上解决目标的尺度变化问题。为了同时考虑目标前景和背景信息,本文给出了一种新的判别跟踪方法,采用多示例学习处理正样本定位模糊问题。该方法是建立在粒子滤波框架下,可以有效解决尺寸变化问题。

2 跟踪算法流程及步骤

本文提出的算法的基本流程如图 1 所示。首先,抽取正负样本并获取它们的外观模型,用一组低维颜色统计特征表示为 $F(X^\alpha)$ 和 $F(X^{\zeta,\beta})$,然后用样本训练贝叶斯分类器。在粒子滤波框架中,根据

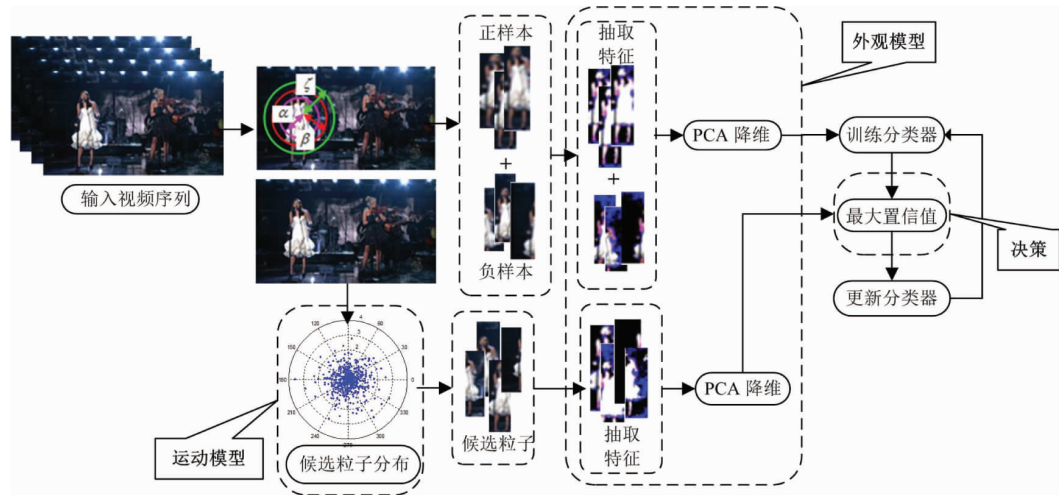


图 1 本文算法流程图

运动模型产生的候选样本来估计置信值。通过分类器 $h_t(\mathbf{x})$ 构建置信函数来估计后验概率 $P(y = 1 | \mathbf{x})$, 其中 $y \in \{0, 1\}$ 是一个代表样本标签的二进制变量。在决策阶段, 根据最大置信值 $\mathbf{x}^* = \arg \max_{\mathbf{x}} (P(y = 1 | \mathbf{x}))$ 判别最佳的候选粒子 $\mathbf{x}^*$ 。

本文采用与在线判别特征选择 (online discriminative feature selection, ODFS)^[10] 跟踪器类似的方式获得正负样本。设 $I_t(\mathbf{x}) \in R^2$ 表示样本 $\mathbf{x}$ 在第 t 帧的位置, 本文中目标位置始终设为 $I_t(\mathbf{x}^*)$ 。抽取两组图像块 $X^\alpha = \{\mathbf{x} | \|I_t(\mathbf{x}) - I_t(\mathbf{x}^*)\| < \alpha\}$ 和 $X^{\zeta, \beta} = \{\mathbf{x} | \zeta < \|I_t(\mathbf{x}) - I_t(\mathbf{x}^*)\| < \beta\} (\alpha < \zeta < \beta)$ 分别表示正样本和负样本。

3 基于颜色统计特征的判别跟踪方法

3.1 颜色统计特征

如何有效地选取特征对跟踪器的性能是很重要

的。Berlin 和 Kay^[17] 提出英语语义中主要包含 11 种基本颜色: 黑色、蓝色、棕色、灰色、绿色、橙色、粉红色、紫色、红色、白色和黄色。为了将色名特征运用到计算机视觉中, Joost 等人^[14] 将颜色命名算法应用于 Berlin 和 Kay 在世界色彩调查 (world color survey, WCS) 中使用的孟塞尔颜色数组中, 并提供了一个映射将 RGB 值与语言颜色标签联系起来, 将 RGB 值映射到一个概率性的 11 维的颜色空间中, 每一维的概率值总和为 1。

然而, 本文使用的是 Martin 等人^[15] 提供的经过标准化处理后的色名映射, 主要是将 11 维的色名特征投影到 10 维子空间的标准正交基中进行标准化, 然后将灰度特征与获取的 10 维色名特征相结合建立最终的颜色统计特征, 通过这种方式获取的特征有效降低了光照的影响。图 2 直观地展示本文中采用的颜色统计特征。第一行显示的是原始图像, 中间一行显示的是抽取颜色统计特征后的特征图像。



图 2 原始图像、颜色特征图像和 PCA 降维重构图像

3.2 颜色统计特征降维

跟踪器的效率与特征维度之间具有显著相关性。虽然颜色统计特征能有效地表示目标外观, 然而, 它的高维度将极大地影响后续的跟踪速度。为了解决这个问题, 本文采用主成分分析 (principal component analysis, PCA) 方法来减少颜色统计特征在每个通道的维度, 既可以去除噪声和背景信息, 同时又保留了有用信息, 大大提高了跟踪速度。

将从图像块中抽取的颜色统计特征向量表示为 $F(\mathbf{v}) = (f_1(\mathbf{v}), \dots, f_L(\mathbf{v}))$, 其中 L 表示颜色统计特征通道数, $f_l(\mathbf{v}) = (v_1, v_2, \dots, v_p) (l = 1, \dots, L)$ 表示第 l 个通道的颜色特征共有 p 维。PCA 将特征向量 $(v_1, v_2, \dots, v_p)$ 整合成一个维数尽可能少的综

合向量 $(x_1, x_2, \dots, x_q) (q < p)$, 新生成的综合向量不仅可以反映原始特征向量的信息, 而且该综合向量的维数之间是相互联系的。

为了确定主成分的数目 q , 通过计算前 q 个主成分的方差贡献率, 即:

$$\frac{\sum_{i=1}^q \lambda_i}{\sum_{j=1}^p \lambda_j} > 0.99 \tag{1}$$

其中, λ_i 和 λ_j 分别表示与特征向量对应的第 i 个和第 j 个特征根。若前 q 个主成分的累积方差贡献率超过 99%, 就用前 q 个主成分 $(x_1, x_2, \dots, x_q) (q < p)$ 的变化来刻画每个特征向量 $(v_1, v_2, \dots, v_p)$ 的变化, 此时特征向量表示为 $F(\mathbf{x}) = (f_1(\mathbf{x}), \dots,$

$f_L(\mathbf{x}))$ 。

为了证明 PCA 降维方法对于本章特征的建立是可行的,本节进行了 PCA 降维重建,重建的结果展示在图 2 最后一行。根据图 2 所示,很明显可以看出,PCA 降维重建的结果和经过颜色统计特征提取后的图像是非常相似的,这证明了经过 PCA 降维后的特征可以代替原来的特征并实现原来特征的功能。

3.3 分类器的构建及更新

在多示例学习(MIL)框架下构建的分类器能够有效地解决正样本的定位模糊问题,同时,为了能更好地适应视觉跟踪中目标外观变化,在线更新目标模型是很有必要的。样本 $\mathbf{x}$ 用特征向量 $\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_L(\mathbf{x}))$ 表示,其中每个特征假定都是独立的, $f_l(\mathbf{v}) (l = 1, \dots, L)$ 代表特征集,分类器 $h_L(\mathbf{x})$ 可用朴素贝叶斯分类器建模如下^[10]:

$$\begin{aligned} h_L(\mathbf{x}) &= \log \left(\frac{\prod_{l=1}^L p(f_l(\mathbf{x}) | y = 1) P(y = 1)}{\prod_{l=1}^L p(f_l(\mathbf{x}) | y = 0) P(y = 0)} \right) \\ &= \sum_{l=1}^L \phi_l(\mathbf{x}) \end{aligned} \quad (2)$$

其中 $\phi_l(\mathbf{x}) = \log \left(\frac{p(f_l(\mathbf{x}) | y = 1)}{p(f_l(\mathbf{x}) | y = 0)} \right)$ 是一个等先验的弱分类器,即 $P(y = 1) = P(y = 0)$ 。由此可见分类器 $h_L(\mathbf{x})$ 是弱分类器 $\phi_l(\mathbf{x})$ 的线性组合。

本文用一组低维的颜色统计特征 $f_l (l = 1, \dots, L)$ 表示样本,设分类器 $h_L(\mathbf{x})$ 中的条件分布 $p(f_l | y = 1)$ 和 $p(f_l | y = 0)$ 服从高斯分布^[10],其中 f_l 是第 l 个低维的颜色统计特征向量,则

$$p(f_l | y = 1) \sim N(\mu_l^+, \Sigma_l^+) \quad (3)$$

$$p(f_l | y = 0) \sim N(\mu_l^-, \Sigma_l^-) \quad (4)$$

上式中的参数 $(\mu_l^+, \Sigma_l^+, \mu_l^-, \Sigma_l^-)$ 可以根据正负实例递归估计如下:

$$\mu_l^+ \leftarrow \eta \mu_l^+ + (1 - \eta) \mu^+ \quad (5)$$

$$\begin{aligned} \Sigma_l^+ &\leftarrow \eta \Sigma_l^+ + (1 - \eta) \Sigma^+ \\ &+ \eta(1 - \eta)(\mu_l^+ - \mu^+)(\mu_l^+ - \mu^+)^T \end{aligned} \quad (6)$$

其中 η 是用于更新的学习率,以缓和相邻帧之间的平衡。 $\mu^+ = \frac{1}{N} \sum_{i=0}^{N-1} f_l(x_i)$, N 是正样本的数量,

$$\Sigma^+ = \frac{1}{N} \sum_{i=0}^{N-1} (f_l(x_i) - \mu^+)(f_l(x_i) - \mu^+)^T$$

和 Σ_l^- 的更新方式与之相同。

在本方法中,假设每个样本特征向量中的每个元素是相互独立的,则

$$\begin{aligned} p(f_l(x) | y = 1) &= p(f_l(x_1) | y = 1) \cdots \\ & p(f_l(x_q) | y = 1) \end{aligned} \quad (7)$$

$$\begin{aligned} p(f_l(x) | y = 0) &= p(f_l(x_1) | y = 0) \cdots \\ & p(f_l(x_q) | y = 0) \end{aligned} \quad (8)$$

其中 $f_l(x_i) (i = 1, \dots, q)$ 表示第 l 个颜色统计特征通道的第 i 个元素, q 表示颜色统计特征的每个通道经过 PCA 降维之后的维数。

3.4 判别跟踪算法

在目标跟踪过程中,相邻帧之间目标移动很小,对新的一帧图像,在上一帧跟踪目标位置周围进行高斯采样,采用仿射变换的 6 个参数为目标从 S_{t-1} 到 S_t 的运动进行建模,状态变量设为 $S_t = (x_t, y_t, \theta_t, s_t, \lambda_t, \varphi_t)$, 其中 $x_t, y_t, \theta_t, s_t, \lambda_t, \varphi_t$ 分别表示目标在第 t 帧 x, y 坐标、旋转角度、尺度变化、宽高比、倾斜度。假设状态变量 S_t 的不同仿射参数之间是相互独立的,并且不同帧之间的状态转移模型服从高斯分布,即 $p(S_t | S_{t-1}) = N(S_t; S_{t-1}, \Sigma)$, 其中 Σ 表示协方差矩阵,它是一个对角阵,对角元素分别对应 6 个仿射参数的方差,即 $\sigma_x^2, \sigma_y^2, \sigma_\theta^2, \sigma_s^2, \sigma_\lambda^2, \sigma_\varphi^2$ 。

算法 1 总结了本文跟踪算法的主要步骤。为了判别最佳粒子,根据分类器 $h_L(\mathbf{x})$ 构建置信函数来估计后验概率密度:

$$c(\mathbf{x}) = P(y = 1 | \mathbf{x}) = \sigma(h_L(\mathbf{x})) \quad (9)$$

其中样本 $\mathbf{x}$ 用特征向量 $\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_L(\mathbf{x}))$ 表示, $y \in \{0, 1\}$ 是二值变量表示样本标签, $\sigma(\cdot)$ 代表 sigmoid 函数,即 $\sigma(z) = (1/(1 + e^{-z}))$ 。最终将具有置信函数值最大的粒子即 $\mathbf{x}^* = \arg\max_{\mathbf{x}}(c(\mathbf{x}))$ 作为跟踪结果。

算法 1: 判别跟踪算法

输入:

- 当前帧 t
 - 第 $t-1$ 帧的跟踪结果 $I_{t-1}(\mathbf{x}^*)$
 - n 个候选粒子 $S_{t-1}^i, i = 1, 2, \dots, n$
-

- 输出:
- 第 t 帧的跟踪结果 $I_t(\mathbf{x}^*)$

1: 抽取正负样本 $\mathbf{X}^\alpha$ 和 $\mathbf{X}^{\zeta,\beta}$;

2: 获取正负样本的外观模型 $\mathbf{F}(\mathbf{X}^\alpha)$ 和 $\mathbf{F}(\mathbf{X}^{\zeta,\beta})$;

3: 构建弱分类器 $\phi_i(\mathbf{x})$ 并根据公式(2)级联为强分类器 $h_L(\mathbf{x})$;

4: 根据 $p(S_i | S_{i-1}) = N(S_i; S_{i-1}, \Sigma)$ 抽取 n 个候选粒子的状态 S_i^i , 并获取它们的外观模型 $\mathbf{F}^i(\mathbf{x})$, 其中 $i = 1, 2, \dots, n$;

5: 根据式(9)计算每个候选粒子的置信值 $c(\mathbf{x})$;

6: 将置信值最大的候选粒子 $\mathbf{x}^* = \operatorname{argmax}_{\mathbf{x}}(c(\mathbf{x}))$ 作为当前帧的跟踪结果 $I_t(\mathbf{x}^*)$;

7: 根据式(5)和(6)更新分类器的参数。

4 实验结果及分析

本节通过大量的实验验证了本文方法的有效性和效率。实验中选用了—个公共数据集^[18]中的 35 个彩色视频序列测试本文跟踪器的性能。进行实验的视频序列具有不同的挑战因素,包括光照变化(IV)、尺寸变化(SV)、变形(DEF)、运动模糊(MB)、快速运动(FM)、面内旋转(IPR)、面外旋转(OPR)、背景杂乱(BC)、遮挡(OCC)、超出视野(OV)以及低分辨率(LR)。同时,将本文设计的跟踪器与 5 种先进的追踪器进行比较,包括跟踪学习检测(TLD)^[5],在线判别特征选择(ODFS)^[10],自适应颜色属性(CN)^[15],自适应局部外观模型(ASLA)^[19]以及稀疏协同模型(SCM)^[20],每个跟踪的实验参数都是从同一测评基准库中获得的^[18]。

4.1 实验配置

本实验在 MATLAB 平台下,配置为 i3 3.3GHz 和 2GB 内存的计算机上实现的。每个测试视频的初始帧位置是通过手动标记的。本实验通过仿射参数将目标区域缩放到 32 × 32 像素数目的图像块,在每一帧中以半径 $\alpha = 4$ 抽取 45 个正样本,抽取负样本集 $\mathbf{X}^{\zeta,\beta}$ 的内外半径分别设为 $\zeta = 8, \beta = 38$ 。弱分类器的学习率设置为 0.93,候选粒子的数目设为 $n = 600$ 。

4.2 定性评价

本节选择了 4 个具有不同挑战因素的视频序

列。图 3 给出了本文提出的方法与一些先进的算法之间的跟踪结果。

图 3(a)展示了篮球运动员在比赛中拦截、运球以及防守等一系列动作。在整个过程中,由于目标外观经历了光照、尺度、遮挡、变形等变化,很多跟踪器不能一直跟踪到最后,漂移到其他球员或者背景上面,但 ODFS、CN 跟踪器以及本文提出的跟踪器能一直跟踪目标。

图 3(b)所示的 bolt 视频序列中,运动员运动剧烈,而且经历了不同程度的面内和面外旋转,CN 跟踪器和本文跟踪器在整个序列中都能够跟踪目标。这个结果表明,颜色统计特征可以有效地阻止跟踪器漂移到背景区域。

在 david 视频序列(图 3(c))中,一个人从黑暗的房间走到明亮的区域。尽管大多数跟踪器能够跟踪到目标,但在目标转身以及尺寸变小时,部分跟踪器不能很好地适应这种变化。ALSA、SCM、TLD、CN 跟踪器和本文方法可以自始至终准确地跟踪移动的目标。由于颜色统计特征能够削弱光照的影响,因此本文跟踪器能使分类器更好地适应目标的外观改变。

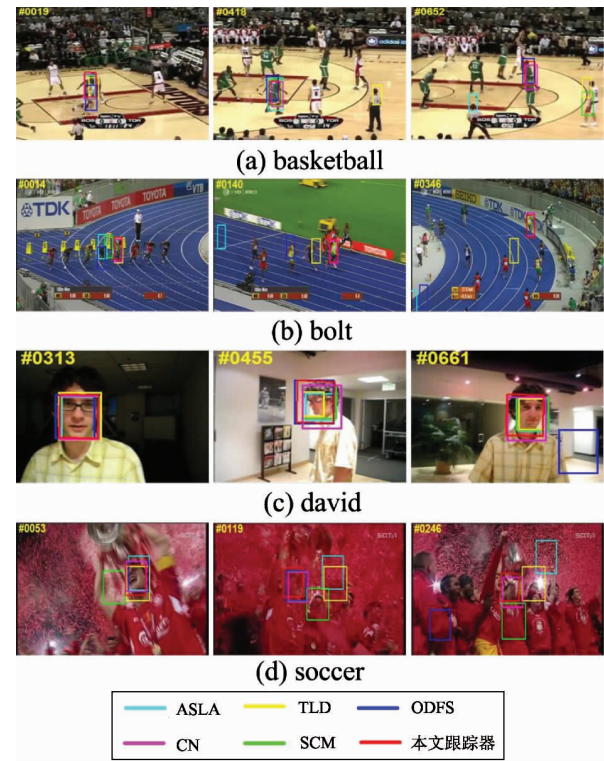


图 3 6 种跟踪器的跟踪结果

Soccer 视频序列展示的是庆祝的场面,背景杂乱,而且背景的颜色和目标衣服颜色很相近,目标在整个过程中经历了遮挡、尺寸变化以及光照变化。由于本文方法有效地利用了多个示例和颜色统计特征更新和训练分类器,因此能很好地处理剧烈的外观变化,很好地跟踪目标对象。

4.3 定量评价

本节采用重叠率(overlap rate, OR)定量评估跟踪器的性能,OR 是通过计算算法跟踪结果目标框与真实结果目标框之间的重叠率。本文用 35 个有挑战性的彩色视频序列进行了测试,表1展示了本文方法和其他 5 种方法对于其中 15 个视频序列的平均重叠率结果,其中最好的两个结果分别用粗体和下划线表示。从表中可以明显地看出,选择的视频序列相对来说具有更多的挑战因素,虽然本文方法的跟踪结果不一定最准确,但相对比较稳定。

图 4 展示了本文算法和 5 种跟踪器针对 35 个测试视频序列的整体性能。为了测试跟踪器是否对

表 1 平均重叠率的比较结果

	ASLA	SCM	TLD	ODFS	CN	本文跟踪器
basketball	0.39	0.47	0.02	0.62	<u>0.64</u>	0.72
bolt	0.01	0.02	0.10	0.01	0.79	<u>0.71</u>
coke	0.17	0.33	<u>0.40</u>	0.23	0.42	0.39
couple	0.08	0.10	0.78	0.58	0.09	<u>0.62</u>
david	0.74	<u>0.72</u>	<u>0.72</u>	0.25	0.54	0.65
david3	0.42	0.39	0.10	<u>0.69</u>	0.61	0.77
deer	0.03	0.07	0.69	0.29	<u>0.76</u>	0.77
football1	<u>0.51</u>	0.42	0.37	0.18	<u>0.51</u>	0.66
lemming	0.15	0.14	0.53	0.63	0.29	<u>0.57</u>
singer1	<u>0.80</u>	0.86	0.73	0.33	0.36	0.78
skating1	0.49	0.46	0.19	0.11	0.55	<u>0.51</u>
subway	0.19	<u>0.69</u>	0.19	0.54	0.19	0.70
tiger1	0.29	0.16	<u>0.50</u>	0.09	0.21	0.55
walking2	0.36	0.78	0.29	0.23	0.32	<u>0.59</u>
woman	0.15	<u>0.67</u>	0.13	0.13	0.20	0.73
Median	0.32	0.42	0.38	0.33	<u>0.43</u>	0.65

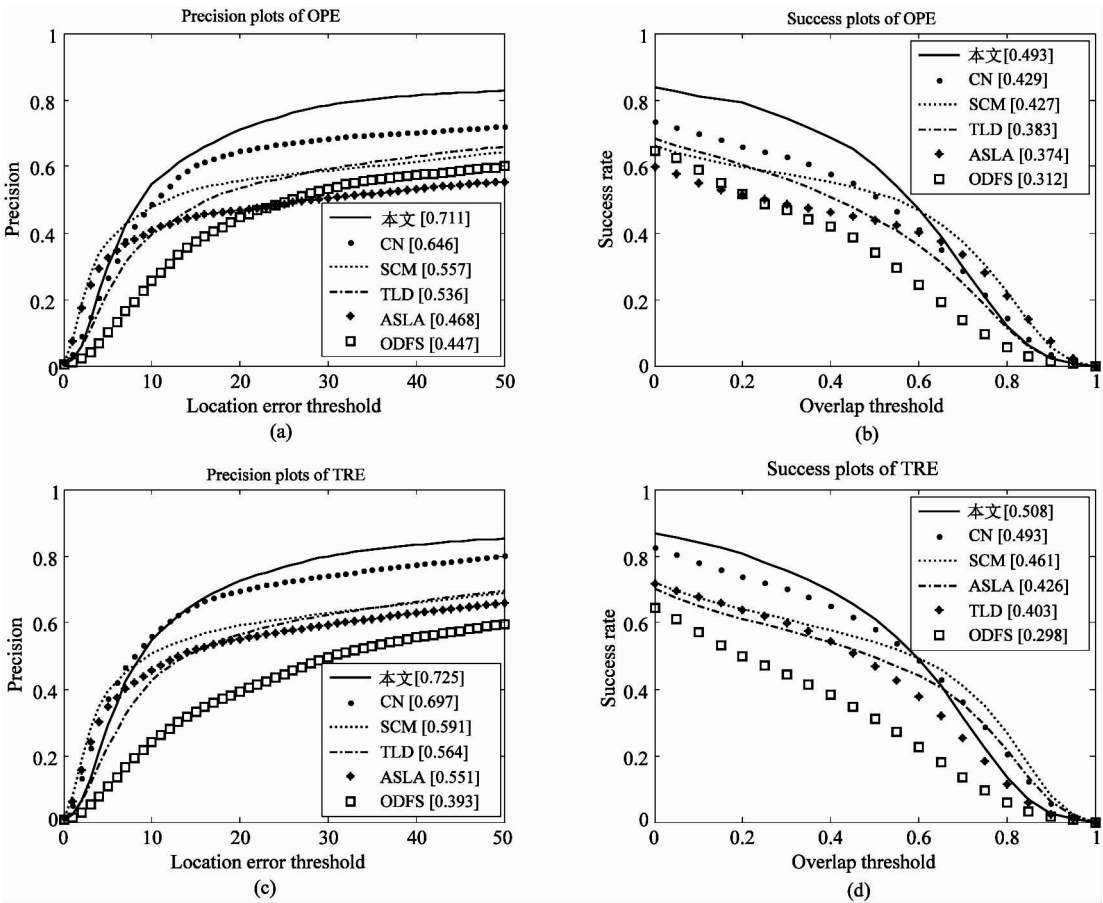


图 4 精确度图和成功率图

初始值敏感,本文采用一次性评估(one-pass evaluation, OPE)和时间鲁棒性评价(temporal robustness evaluation, TRE)^[18]来评估本文提出的跟踪器的整体性能。OPE 通过将跟踪器从视频序列的第一帧开始运行,获得平均精确度(precision)和成功率(success rate)。TRE 则不同,采用的是将跟踪器从视频的不同帧开始运行到最后。

图 4(a)(c)评估的是算法的精度,其中横坐标表示位置误差阈值(location error threshold),单位是像素,纵坐标表示精度。位置误差是通过计算算法跟踪结果与真实结果的中心位置的欧式距离得到的,若位置误差小于选定的阈值,则认为跟踪结果准确。从图 4(a)(c)可以看出,选定的位置误差阈值越大,同一个跟踪器的精度也越大。本实验选定误差阈值为 20 像素时对应的精度估测不同的跟踪器,精确度越高说明跟踪器性能越好。相应的,图 4(b)(d)评估的则是算法的成功率,其中横坐标表示重叠阈值(overlap threshold),纵坐标代表成功率。重叠率的计算方式和 OR 相同,当重叠率大于选定

的阈值时,则认为跟踪结果准确。因此,从图 4(b)(d)中可以看出,当重叠阈值越大,同一个跟踪器的成功率反而越小。在这两个图中,采用曲线下面积(AUC)来评估不同的跟踪器,AUC 越大的跟踪器性能越好。从图 4 可以明显看出,无论是 OPE 还是 TRE,本文提出的方法性能都比较好,这说明本文方法对初始值不敏感,但其他跟踪器在这 4 个图中的排名却并不一致,这是因为有的跟踪器对初始值相对比较敏感。

跟踪器的性能会受不同挑战因素的影响。图 5 举例说明和分析了本文跟踪器在变形(DEF)、遮挡(OCC)以及尺寸变化(SV)这三种挑战因素下的成功率图和精确度图。从图 5 中可以看出,本文跟踪器的性能是比较优异的,除了在图 5(c)中成功率比 SCM 跟踪器低了 1.4% 位居第二。除了图 5(c),本文跟踪器和 CN 跟踪器性能都超过其他跟踪器,这显示了颜色统计特征在目标跟踪中的优越性。此外,由于采用了仿射变换,相对于 CN 跟踪器来说,本文的方法能更好地适应目标尺度的变化。

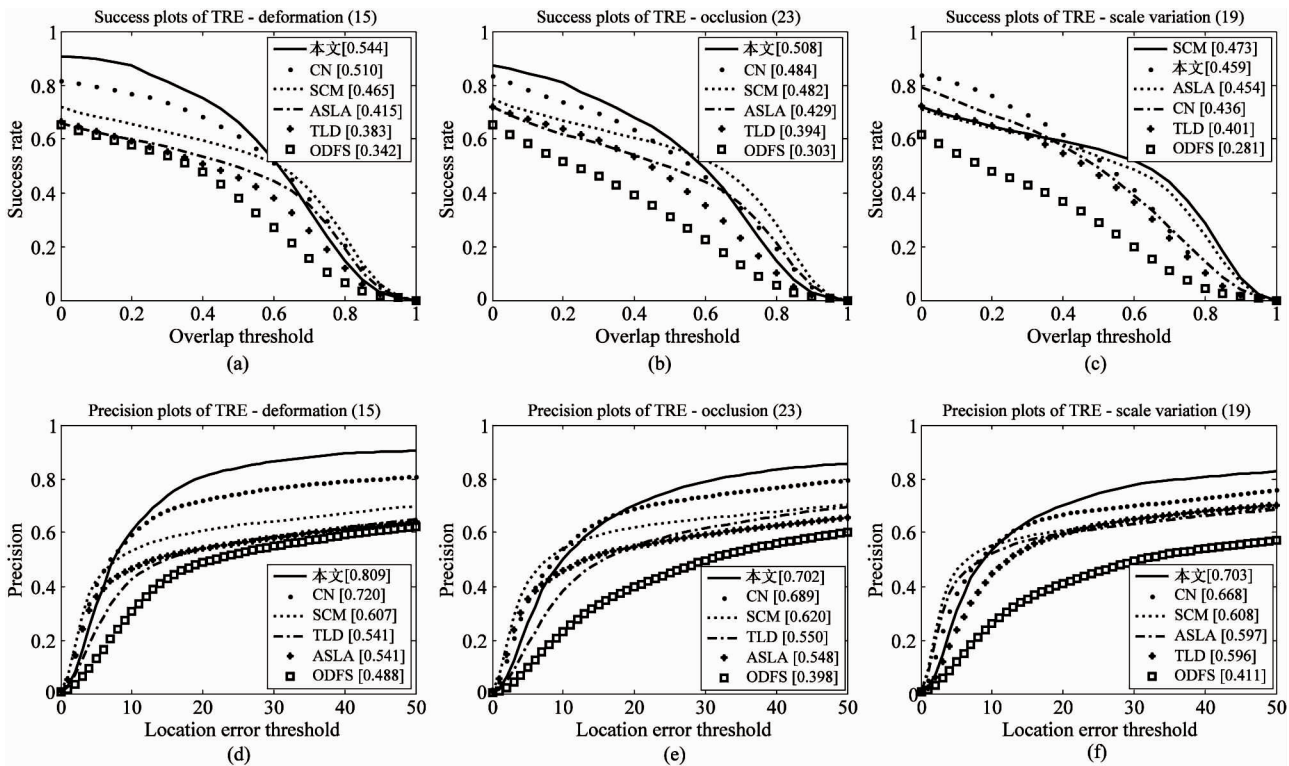


图 5 在变形(DEF)、遮挡(OCC)以及尺寸变化(SV)等不同挑战因素下的精确度图和成功率图

5 结 论

本文提出了一个新的简单而鲁棒的基于颜色统计特征的判别跟踪算法。本文在跟踪过程中建立了仿射运动模型和优化参数来解决尺寸变化以及角度变化等问题。此外,采用低维特征空间表示保存有用的颜色信息来反映目标外观变化,并直接用颜色统计特征训练分类器,避免了特征选择,从而设计了一个更简单的方法。大量的实验结果表明,在面对不同的挑战因素时,本文提出的跟踪方法都有较大的优势。然而,本文方法在提高准确性的同时降低了帧率。今后将会针对这一问题,寻求解决的方案以使跟踪器的精确度和效率更高。

参考文献

- [1] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram. In: IEEE Conference on Computer Vision & Pattern Recognition, New York, USA, 2006. 798-805
- [2] Kwon J, Lee K M. Visual tracking decomposition. In: IEEE Conference on Computer Vision & Pattern Recognition, San Francisco, USA, 2010. 1269-1276
- [3] Ross D A, Lim J, Lin R S, et al. Incremental learning for robust visual tracking. *Advances in Neural Information Processing Systems*, 2008, 77(1-3): 125-141
- [4] 郭鹏宇, 苏昂, 张红良等. 结合纹理和形状特征的在线混合随机朴素贝叶斯视觉跟踪器. *光学学报*, 2015, 03: 195-205
- [5] Kalal Z, Matas J, Mikolajczyk K. P-N learning: Bootstrapping binary classifiers by structural constraints. In: IEEE Conference on Computer Vision & Pattern Recognition, San Francisco, USA, 2010, 238(6): 49-56
- [6] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: IEEE Conference on Computer Vision & Pattern Recognition, San Diego, USA, 2005. 886-893
- [7] Ahonen T, Hadid A, Pietikäinen M. Face recognition with local binary patterns. *Lecture Notes in Computer Science*, 2004, 3021: 469-481
- [8] 朱秋平, 颜佳, 张虎等. 基于压缩感知的多特征实时跟踪. *光学精密工程*, 2013, 21(2): 437-444
- [9] Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2011, 33(8): 1619-1632
- [10] Zhang K, Zhang L, Yang M H. Real-time object tracking via online discriminative feature selection. *IEEE Transactions on Image Processing*, 2013, 22(12): 4664 - 4677
- [11] Swain M J, Ballard D H. Color indexing. *International Journal of Computer Vision*, 1991, 7(1): 11-32
- [12] Stricker A M A, Orengo M. Similarity of color images. *Proceedings of SPIE - The International Society for Optical Engineering*, 1970, 2420: 381-392
- [13] Huang J, Kumar S R, Mitra M, et al. Image indexing using color correlograms. In: IEEE Conference on Computer Vision & Pattern Recognition, San Juan, Puerto Rico, 1997. 762-768
- [14] Joost V D W, Cordelia S, Jakob V, et al. Learning color names for real-world applications. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2009, 18(7): 1512-1523
- [15] Danelljan M, Khan F S, Felsberg M, et al. Adaptive color attributes for real-time visual tracking. In: IEEE Conference on Computer Vision & Pattern Recognition, Columbus, USA, 2014. 1090-1097
- [16] Avidan S. Ensemble tracking. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2007, 29(2): 261-271
- [17] Berlin B, Kay P. Basic Color Terms: Their Universality and Evolution. Berkeley: UC Press, 1969
- [18] Wu Y, Lim J, Yang M H. Online object tracking: A benchmark. In: IEEE Conference on Computer Vision & Pattern Recognition, Portland, USA, 2013. 2411-2418
- [19] Jia X. Visual tracking via adaptive structural local sparse appearance model. In: IEEE Conference on Computer Vision & Pattern Recognition, Providence. USA, 2012. 1822-1829
- [20] Zhong W. Robust object tracking via sparsity-based collaborative model. In: IEEE Conference on Computer Vision & Pattern Recognition, Providence, USA, 2012. 1838-1845

A new discriminative tracking algorithm based on color statistical characteristics

Hu Zhaohua^{* **}, Ju Rong^{*}, Ou Yangwen^{*}, Jin Rong^{*}

(^{*} School of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044)

(^{**} Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing 210044)

Abstract

In view of the problem that most discriminative algorithms for target tracking focus on the tracking speed while neglect the scale variation, a simple, robust discriminative algorithm based on color statistical characteristics was presented. The new color statistical characteristics not only possess the certain illumination invariance, but also maintain the higher discriminative power. An affine kinematics model for tracking was established to keep optimizing the parameters during tracking to solve the scale variation and the view angle change. To improve the tracking speed, the low-dimensional color statistical characteristics were used to describe the target appearance, and the color statistical characteristics were used to train naive Bayes classifiers and update classifiers online. The sample with the maximum confidence was regarded as the tracking result. Numerous comprehensive experiments were conducted for evaluation of the proposed algorithm and other algorithms, and the remarkable effectiveness under different challenge factors of the proposed tracking algorithm was showed.

Key words: visual tracking, color statistical characteristics, affine transformation, dimensionality reduction, discriminative classifier