

融合 SURF 特征的压缩追踪算法

方露, 韩超

(安徽工程大学电气工程学院, 安徽 芜湖 241000)

摘要:针对压缩追踪(Compressive Tracking, CT)算法在目标追踪中易受遮挡和扭曲变形干扰问题,结合该算法简单容易执行的追踪机制,提出一种融合 SURF(Speeded-up robust features)和压缩特征的鲁棒性目标追踪算法。新算法有两方面的改进:一是在自适应更新目标外观模型的基础上,增加防止误更新外观模型机制,解决追踪过程中严重遮挡和扭曲变形问题;二是通过 SURF 特征点在前后两帧中的匹配关系,求解追踪目标尺寸变化,自适应调整目标模板大小。通过仿真实验表明:改进后的算法在公开的某些图像序列上的追踪效果良好,与 CT 算法及改进的 CT 算法相比正确性和鲁棒性上性能更优越。

关键词:压缩追踪算法;SURF 算法;误更新机制;追踪模板

中图分类号:TP181

文献标志码:A

引言

视频追踪在日常生活中应用广泛,它在安全监控、导航定位、智能仪器上有着极其重要的作用,虽然目前有很多算法对视频中的物体进行追踪,但是由于目标在长时间运动过程中会面临各种不确定的干扰,如其他物体的遮挡、自身的形变、环境明暗变化等^[1],因此提出一种鲁棒的、高效的目标追踪算法显得尤为重要。

很多学者对物体追踪问题做出了许多突出的贡献。文献[2]提出了一种高效的基于压缩感知理论的物体追踪算法,测试结果表明,这是一种高效的、鲁棒的追踪算法,但该算法在针对遮挡和复杂背景的情况下存在缺陷。文献[3]提出一种改进的实时压缩追踪算法,该算法对遮挡、尺度变化等问题进行处理,改进后的压缩追踪算法鲁棒性和跟踪速度得到改善。文献[4]提出一种融合 SIFT 特征的压缩追踪算法,该算法在文献[3]的基础上做进一步的完善,实验结果表明比文献[3]的追踪

效果更加有效,但该算法采用的 SIFT 进行特征点检测耗时较长,实现目标实时性追踪要求较高。文献[5]对原始追踪目标进行特征提取,利用空间投影的方法得到两种特征,这种改进的压缩感知追踪算法实现两种特征的融合,相对单一的特征鲁棒性更好,但该算法的计算复杂度高。文献[6]在 Mean Shift 算法的基础上融合 SURF 特征,提出一种自适应分块目标追踪算法,该算法虽然提高了追踪精度,但在遮挡或光照变化的情况下仍然存在漂移的情况。文献[7]利用特征选择与目标模型更新结合的方法进行压缩追踪,实现算法鲁棒性增强和追踪速度加快功能,该算法在特征选择与目标模型的结合方面不是很完善。

总结以上学者的工作,提出一种融合 SURF 特征的压缩追踪算法。该算法改进的方面包括两点:一是增加误更新外观模型机制,有效地解决了遮挡和扭曲变形问题;二是采用 SURF 特征点匹配,完成追踪模板的自适应调整。

收稿日期:2016-04-15

基金项目:安徽省自然科学基金资助项目(1508085MF121);安徽高校自然科学基金项目(KJ2016A056);安徽检测技术与节能装置省级实验室开放研究基金(1506c085002);高校优秀中青年骨干人才国内外访学研修重点项目(gxfxZD2016100)

作者简介:方露(1991-),男,安徽舒城人,硕士生,主要从事目标检测、视频追踪方面的研究,(E-mail)ahpu3008@163.com;

韩超(1974-),男,安徽宿州人,副教授,博士,主要从事图像处理、计算全息及全息显示等方面的研究,(E-mail)hanchaozh@126.com

1 压缩感知追踪算法简介

1.1 压缩特征提取

对每一个样本 $z \in \mathcal{R}^{w \times h}$, 通过将其与一系列多尺度的矩阵滤波器 $\{h_{1,1}, \dots, h_{w,h}\}$ 进行卷积, 矩形滤波器定义为:

$$h_{i,j}(x,y) = \begin{cases} 1, & 1 \leq x \leq i, 1 \leq y \leq j \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

式中, i 和 j 分别是矩形滤波器的宽和高。滤波后的图像矩阵由 $w \times h$ 维的列向量表示, 则列向量表示为 $z' \in \mathcal{R}^{(w \times h) \times l}$, 将这些列向量连接成一个高维的多尺度图像向量 $x = (x_1, \dots, x_m) \in \mathcal{R}^{(w \times h)^2 \times l}$, 维数高达 $10^6 \sim 10^{10}$, 由此可见该矩阵计算量过大, 达不到实时性要求。为此, 文献[2]提出了一种稀疏的随机测量矩阵, 该矩阵满足受限等距性质(RIP)条件^[8], 矩阵中的元素 r_{ij} 满足公式:

$$r_{ij} = \sqrt{s} \times \begin{cases} 1 & p = \frac{1}{2s} \\ 0 & p = 1 - \frac{1}{s} \\ -1 & p = \frac{1}{2s} \end{cases} \quad (2)$$

由文献[2]可知: p 为概率, $s = (w \times h)^2/4$ 且 s 的值在 2 ~ 4 之间随机选取。当 $s = 3$ 时, 计算的数据量只有原来的 1/3。

由压缩感知理论可知: 对于 $n \times m$ 维 ($n \ll m$) 的随机矩阵 $R \in \mathcal{R}^{m \times n}$, 将高维图像空间的 $x \in \mathcal{R}^m$ 变换到一个低维的空间 $v \in \mathcal{R}^n$, 数学表达式为:

$$v = Rx \quad (3)$$

低维的 v 理想情况下保留高维的 x 的信息, 或者保持原始空间中各样本 x 的距离关系^[9]。由此可知, 压缩后的特征是原始图像区域特征加权和, 压缩特征提取示意图如图 1 所示。

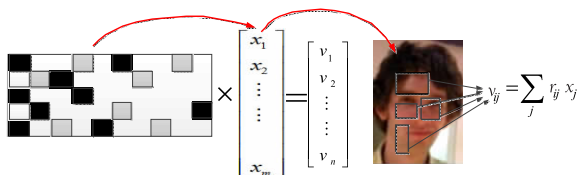


图 1 压缩特征提取示意图

1.2 训练朴素贝叶斯分类器

对每个样本 $z \in \mathcal{R}^m$ (m 维向量), 它的低维表示是 $v = (v_1, \dots, v_n)^T \in \mathcal{R}^n$ (其中 $n \ll m$)。假定 v 中的各元素是独立分布的, 可以通过朴素贝叶斯分类器来建立模型。

$$H(v) = \log \left(\frac{\prod_{i=1}^n p(v_i | y = 1) p(y = 1)}{\prod_{i=1}^n p(v_i | y = 0) p(y = 0)} \right) = \sum_{i=1}^n \log \left(\frac{p(v_i | y = 1)}{p(v_i | y = 0)} \right) \quad (4)$$

其中: y 表示样本类别, y 为 1 时或 0 时分别表示正、负样本。分类器中的条件分布 $p(v_i | y = 1)$ 和 $p(v_i | y = 0)$ 符合高斯分布:

$$\begin{cases} p(v_i | y = 1) \sim N(\mu_i^1, \sigma_i^1) \\ p(v_i | y = 0) \sim N(\mu_i^0, \sigma_i^0) \end{cases} \quad (5)$$

式中的 μ 和 σ 分别为样本的均值和标准差。

1.3 分类器参数更新

样本的均值和标准差的计算公式为:

$$\begin{cases} \mu^t = \frac{1}{n} \sum_{k=0}^{n-1} v_i(k) \\ \sigma^t = \sqrt{\frac{1}{n} \sum_{k=0}^{n-1} (v_i(k) - \mu^t)^2} \end{cases} \quad (6)$$

其中, 第 k 个正样本数为 $v_i(k)$, n 为正样本个数。分类器参数更新公式为:

$$\begin{cases} \mu_i^1 \leftarrow \lambda \mu_i^1 + (1 - \lambda) \mu^1 \\ \sigma_i^1 \leftarrow \sqrt{\lambda (\sigma_i^1)^2 + \lambda (1 - \lambda) (\mu_i^1 - \mu^1)^2} \end{cases} \quad (7)$$

按照同样的方式对 (μ_i^0, σ_i^0) 进行更新, λ 为学习率, 决定着更新速度的快慢。

2 改进追踪算法

2.1 目标模型更新方法

在基于检测的压缩追踪过程中, 通常将目标与背景作为一个二分类问题来看待, 在文献[2]中根据分类器的响应值较大者作为新的跟踪结果。响应值公式为:

$$\sigma(H(v)) = \frac{1}{1 + e^{-H(v)}} \quad (8)$$

对于一个二分类问题, 对追踪目标模型的更新显得十分重要, 当 $H(v)$ 大于零, 由公式(8)可知函数值大于 0.5, 结合图 2 可知, 此时根据响应结果判断出样本为正, 其他情况视为负样本。

在追踪目标的过程中, 伴随着环境的变换和目标自身因素的影响, 噪声的引入是不可避免的, 当目标遭遇严重遮挡时或者姿态变化频繁, 此时采集的样本会出现响应值全部小于零的情况, 这种检测结果会错误地更新目标模型, 导致后面的追踪失败。因此, 本文针对目标样本的分类器响应最大值小于零的情况, 提出分类器不予更新思想。得到的公式化的结果为:

$$H_{\max} = \max(H(x_i)) = \begin{cases} < 0, \text{不更新} \\ \geq 0, \text{更新} \end{cases} \quad (9)$$

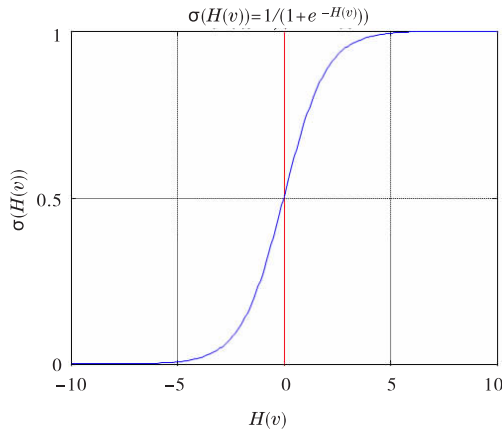


图 2 sigmoid 函数示意图

2.2 SURF 特征融合

由于压缩追踪算法不能实时的改变目标追踪窗口的大小,随着目标运动状态的改变,伴随着环境的改变,目标大小的变化影响着正负样本的采集,导致长时间追踪过程中存在漂移的情况。文献[4]中提出基于 SIFT 特征点提取的压缩追踪算法,SIFT 算法在局部区域检测和匹配性能方面表现出良好的效果,但是由于其自身的计算复杂度高,在对目标长时间追踪过程中,实时性要求较高。针对该算法在处理目标尺寸方面存在问题,结合文献[4]在实时性方面的不足,提出一种融合 SURF 特征的压缩追踪算法,SURF 算法计算速度快,在实时性方面有很大的优势,同时 SURF 的尺度不变特性^[10]在应对尺度缩放、旋转和亮度变化方面具有不变性的优势。

该算法能够根据运动目标的变化,对前后两帧利用 SURF 算法进行特征点的提取,通过仿射变换计算得到仿射变换矩阵。相邻两帧中的目标一般满足的仿射变换关系如下公式所示:

$$I_{back} = I_{first} \cdot M \quad (10)$$

式中, I_{back} 和 I_{first} 分别表示后面一帧和前面一帧图像, M 为仿射矩阵。

$$M = \begin{bmatrix} \lambda \cos\theta & \lambda \sin\theta & 0 \\ \lambda - \sin\theta & \lambda \cos\theta & 0 \\ x & y & 1 \end{bmatrix} \quad (11)$$

其中: λ 是尺度因子, θ 是旋转因子, (t_x, t_y) 是平移向量。对于参数的求解,使用随机抽样一致性 RANSAC 模型来解决,SURF 特征点匹配结果如图 3 所示。

2.3 新算法流程

根据初始帧目标位置和尺度信息,生成随机稀疏矩阵,对目标图像进行压缩采样,并训练生成初始的贝叶

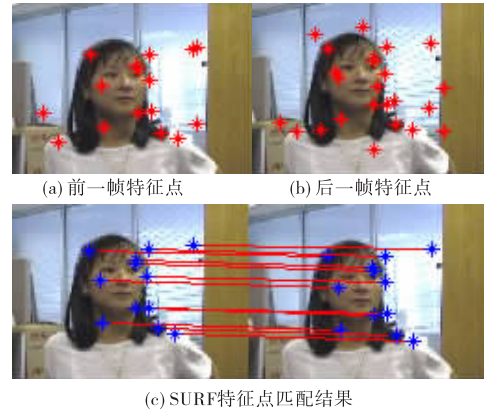


图 3 SURF 匹配结果

斯分类器。后续的每帧图像通过新算法在线执行五个步骤:

(1) 在第 t 帧图像,给定搜索半径 γ , 以上一帧目标位置 l_{t-1} 为基准,采样得到的样本位置集合为 $D^\gamma = \{z | l(z) - l_{t-1} < \gamma\}$ 。

(2) 通过前面公式提取出对应的低维特征 $V(z) = (v_1, \dots, v_n)^T$, 用式(4)计算样本集合 D^γ 中低维特征分类器响应值 $H(v)$ 。

(3) 其中分类器响应值最大的即为当前帧的目标位置 l_t , 作为新的跟踪结果,如果 $H_{\max} < 0$, 不更新分类器,转至步骤(1)进行下一帧处理,否则进入步骤(4)。

(4) 使用 SURF 特征调整目标尺度和旋转尺度,根据公式(11),利用 RANSAC 模型求解参数,获得目标尺度,进行自适应调整。

(5) 以当前帧的目标位置为 l_t 基准,给定搜索半径 α, β 和 ζ , 采样得到正样本集合和负样本集合分别为 $D^\alpha = \{z | l(z) - l_t < \alpha\}$ 、 $D^{\zeta, \beta} = \{\zeta < z | l(z) - l_t < \beta\}$, 再通过式(3)提取两个集合中样本对应的低维特征,利用式(7)对分类器参数进行更新。

3 实验结果与分析

3.1 实验条件

为了验证本文提出算法的有效性、合理性,将本文算法与经典的 CT 算法和文献[4]改进的 CT 算法(下文简称文献[4]算法)进行比较和分析。采用了三组代表性的视频序列作为测试样本,分别是有关遮挡的 Girl 视频序列,快速变化的 Jumping 视频序列,光照变化的 Trellis 视频序列,所有样本均来源于标准的视频库。所有算法在 Windows7 系统下 MATLAB R2012b 运行,计算机配置是 AMD A6-3400M APU, 3 GB RAM。三种算法对 Girl、Jumping、Trellis 标准视频的部分视频追踪结果如图 4~图 6 所示。



图4 Girl(第20、70、140、171、260、456帧)

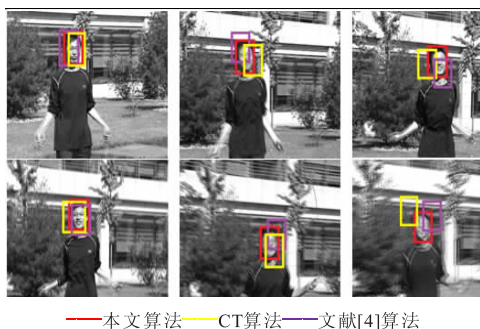


图5 Jumping(第23、70、131、175、275、306帧)



图6 Trellis(第19、80、150、290、419、550帧)

3.2 实验结果定性分析

如图4所示,分别是对 Girl 序列中第 20、70、140、171、260、456 帧的追踪结果。在 Girl 视频序列中选择追踪的目标为女孩的脸,在前面 150 帧内,三种算法在目标追踪的效果上差别不大,原因是目标发生变化不大,背景环境的干扰较小。当 150 帧以后,传统的 CT 算法针对目标自身的移动发生轻微的漂移现象,但其他两种算法的追踪效果相差不大。400 帧以后由于出现遮挡的情况,本文和文献[4]的算法目标追踪模板发生自适应更新,CT 算法追踪模板没有改变,随着追踪的进行,目标被完全的遮挡,此时文献[4]算法和 CT 算法开始出现错误的更新目标模型,出现了严重的漂移现象。随着后续的进行,以及正负样本的更新,三种算法的追踪效果也趋于稳定。

如图5所示,分别是对 Jumping 序列中第 23、70、131、175、275、306 帧的追踪结果。在 Jumping 视频序列中选择追踪的目标依然是人脸,由于追踪目标快速的变

化,目标特征需要实时更新,CT 算法在目标发生快速移动的过程中,错误的更新特征,导致跟踪效果较差。文献[4]算法虽然采用了追踪目标窗口自适应的改变,融合了 SIFT 特征后相对 CT 算法单一特征追踪效果较好,但是目标长时间的快速变化导致最后目标模型更新错误。而本文采用 SURF 和目标模型更新策略相结合的方式,在应对目标快速变化的过程中表现出更好的追踪效果。

如图6所示,分别是对 Trellis 序列中的第 19、80、150、290、419、550 帧的追踪结果。在 Trellis 视序列中,在前面 140 帧内,由于追踪目标所处的环境光照变化不大,三种算法均能很好地追踪到目标。随着目标的移动,光照变化缓慢,CT 算法和文献[4]算法出现了轻微的漂移,没有丢失目标,本文算法仍然可以准确地追踪到目标。200 帧以后,由于光照剧烈的变化,三种算法都出现了漂移现象,说明改进的算法在应对光照变化的情况下,追踪效果不是很理想,这是后面继续研究中有待改进的方面。

3.3 实验结果定量分析

本文采用中心位置定位误差准则进行评价,该准则采取目标追踪矩形框的质心与真实矩形框质心之间的欧式距离作为评判标准^[11],对三个视频中心误差的测试结果如图7~图9所示。

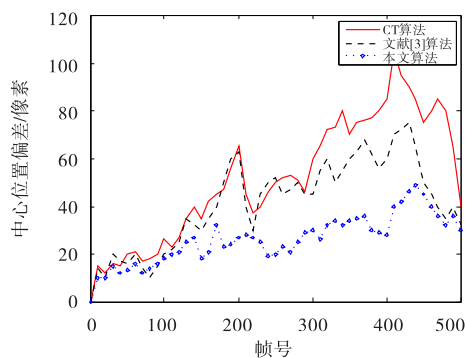


图7 Girl 视频中心位置定位误差分布图

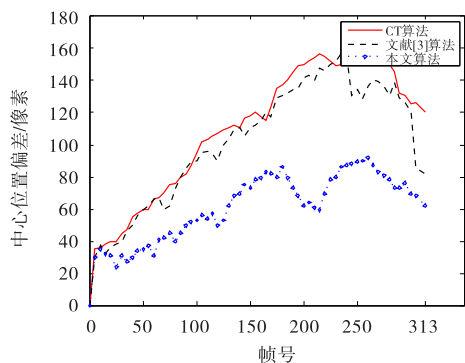


图8 Jumping 视频中心位置定位误差分布图

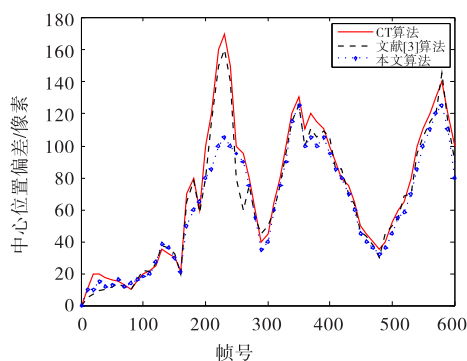


图 9 Trellis 视频中心位置定位误差分布图

4 结束语

本文提出一种融合 SURF 特征的压缩追踪算法,该算法在经典的 CT 算法基础上做进一步的改进。从第一帧图像中提取压缩特征,采集正负样本训练朴素贝叶斯分类器,针对 CT 算法面对目标遮挡和剧烈变化存在的漂移情况,改进的追踪算法在前后两帧目标检测的过程中,采用 SURF 特征点匹配,实现追踪目标窗口大小自适应改变。此方法可以加快追踪速率降低正负样本采集带来的干扰,面对严重遮挡和剧烈变化问题,本文采用一种新的目标模型更新方法,即在目标样本的分类器响应最大值小于零的情况,提出分类器不予更新思想,改善了因为误更新目标模型带来的误差。本文改进后的算法与经典的 CT 算法和文献[4]算法进行定性和定量的分析比较,实验结果表明本文算法在目标严重遮挡和剧烈变形情况下追踪效果优于其他两种,然而,在面对光照变化的问题上本文提出的算法效果不是十分理想,这是后期研究重点方向之一。

Compressive Tracking Algorithm Based on SURF Feature Fusion

FANG Lu, HAN Chao

(College of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, China)

Abstract: Aim at the compressive tracking (CT) algorithm which easily to be occluded and distorted in target tracking, combining the simple tracking mechanism and easy implement, a robust target tracking algorithm which includes SURF (Speeded-up robust features) and compression features is proposed. The new algorithm has two aspects of improvement as following: Firstly, based on the adaptive update of the target appearance model, the mechanism of the appearance model is added to prevent the false appearance, which solves the problem of serious occlusion and distortion in the tracking process. Secondly, through the matching relationship of the SURF feature points between adjacent frames, the solution of the target size change is completed and the adaptive adjustment of tracking template size is achieved. The simulation results show that the improved algorithm has a good tracking effect on some image sequences, and it is superior in accuracy and robustness in comparison with the CT algorithm and the improved CT algorithm.

Key words: compressive tracking algorithm; SURF algorithm; error update mechanism; tracking template

参考文献:

- [1] LI X, HU W M, SHEN C H, et al. A survey of appearance models in visual object tracking[J]. ACM Transactions on Intelligent Systems and Technology, 2013, 4(4): 5801-5848.
- [2] ZHANG K H, ZHANG L, YANG M H. Real-time compressive tracking [C]//Proceedings of Europe Conference on Computer Vision, Florence, Italy, October 7-13, 2012: 864-877.
- [3] 钟权, 周进, 吴钦章, 等. 一种改进的实时压缩跟踪算法[J]. 光电工程, 2014, 41(4): 1-8.
- [4] 钟权, 周进, 崔雄文. 融合 SIFT 特征的压缩跟踪算法[J]. 光电工程, 2015, 42(2): 66-72.
- [5] 唐宇, 凌志刚, 李建成, 等. 基于自适应特征融合的压缩感知跟踪算法[J]. 计算机工程与应用, 2015, 51(10): 160-165.
- [6] 刘振兴, 范新南, 李敏. 融合 SURF 特征的改进自适应分块目标跟踪算法[J]. 计算机工程与设计, 2016, 37(2): 454-459.
- [7] 石武祯, 宁纪锋, 颜永丰. 压缩感知跟踪中的特征选择与目标模型更新[J]. 中国图象图形学报, 2014, 19(6): 932-939.
- [8] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [9] CANDÈS E, TAO T. Near optimal signal recovery from random projections and universal encoding strategies[J]. IEEE Transactions on Information Theory, 2006, 52(12): 5406-5425.
- [10] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF)[J]. Computer Vision & Image Understanding, 2008, 110(3): 346-359.
- [11] 李俊, 谢维信, 李良群. 基于空时线索的 TLD 视频跟踪算法[J]. 信号处理, 2015, 31(10): 1287-1293.