

doi: 10.7690/bgzdh.2014.03.017

基于改进 Haar-like 特征的压缩跟踪算法

袁建建, 毛征, 曲劲松, 吴珍荣, 李红岩

(北京工业大学电子信息与控制工程学院, 北京 100124)

摘要: 针对原始算法特征可能出现的特征无法准确表达目标特性的问题, 提出一种改进 Haar-like 特征的压缩跟踪算法。原始算法利用正负样本训练构造分类器, 利用分类器对候选样本判定, 得到最高分类器响应样本就是目标。进行重采样以更新分类器为下一帧做准备, 对出现的问题, 使用了一种新的图像特征来表示目标特性, 同时加入一系列策略处理样本, 去除那些与目标差异较大的样本, 并进行仿真。仿真结果表明: 该算法不仅提高了分类器对于正负样本的判别性, 也降低了算法的计算复杂度, 提高了算法的实时性。

关键词: 目标跟踪; 压缩感知; Haar 特征; 实时跟踪

中图分类号: TP391 **文献标志码:** A

Compressive Tracking Method Based on Improved Haar-like Feature

Yuan Jianjian, Mao Zheng, Qu Jinsong, Wu Zhenrong, Li Hongyan

(College of Electronic Information & Control Engineering, Beijing University of Technology, Beijing 100124, China)

Abstract: Real-time robust tracking problem is a great challenge in the tracking area. A compressive tracking method based on improved Haar-like feature is proposed. The original method utilizes the positive and negative samples to train a classifier, then the classifier is used to discriminate the candidate samples. The candidate sample which gets the highest classify score is the target. After that the resample are utilized to update the classifier to get ready for next frame. However the original method has some problems. First, the features selected have too much randomness, so the target cannot be well represented by the features selected in the initial stage. Second, all candidate samples are decided by the classifier which need too much calculation, this will affect the real-time quality. To these problems, this paper uses a new image feature to represent the target and embed some methods to pre-process the samples to remove the samples which have little similarity with the target. This can increase the discriminate power of the classifier and decrease the computational complexity which improves the real-time quality of the method.

Keywords: target tracking; compressive sensing; Haar feature; real-time tracking

0 引言

由于环境复杂、光照、物体形变、物体旋转等一系列原因, 长时间稳定的跟踪仍然是相当有难度的工作。基于外表模型可将跟踪算法分为 2 类: 基于生成模型和基于判别模型。

生成模型是通过大量样本进行学习获得目标模型, 利用该模型寻找最小重构误差的图像区域来达到跟踪目的。IVT 算法是一种增量学习算法, 利用增量空间方法来自适应解决目标的形态变化, 在环境变化情况下仍然有很好的效果^[1]。 l_1 -跟踪器认为每块图像纹理是若干个纹理模板的稀疏表示, 通过求解 l_1 最小问题, 来寻找和纹理模板相似度最高的目标^[2]。但是 l_1 -跟踪器计算量太大, 导致实时性能比较差。尽管在线学习的生成模型获得不错的成果, 但是还有一些问题需要解决: 1) 生成模型的在线学习阶段需要大量的采样, 然而一般跟踪开始阶段的样本不会太多。大多数文献都假设目标在开始

阶段的形态不会有太大变化, 这样的假设过于理想化。一旦目标形态发生变化, 导致得到的生成模型在跟踪过程出现跟踪漂移问题。2) 只利用正样本学习构建目标外表模型, 可能会因为某些错误正样本的影响而导致跟踪器偏离正确目标位置, 导致漂移问题出现。3) 这些生成模型的构建只利用正样本信息, 舍弃目标的背景信息, 这样构建的分类器对于背景和目标的区分性不够好。

判别模型是将跟踪问题看作二元分类问题, 通过得到目标和背景的判定边界来区分目标和背景。Avidan 将支持向量机融入光流法以扩展光流法进行目标跟踪^[3]。Collins 等人认为跟踪效果主要取决于跟踪特征对目标和背景的区分能力, 而最具区分能力的特征应该通过在线学习时选择, 以达到准确稳定的跟踪^[4]。然而这些跟踪器只利用正样本和很少的负样本来更新分类器。当利用错误或噪声比较多的样本进行更新目标模型时, 会导致漂移问题。

收稿日期: 2013-10-20; 修回日期: 2013-11-13

作者简介: 袁建建(1986—), 男, 山东人, 硕士, 从事图像处理、视觉跟踪研究。

因此笔者提出了一种改进的基于压缩特征的跟踪算法，增强分类器对样本的判别性，增加系统鲁棒性。在跟踪阶段减少待判定样本的数目而降低算法计算复杂度，提高算法实时性。

1 相关算法

压缩感知理论指出信号在某一变换域是稀疏的或者说是可压缩的，那么可以设计一个与相变换基不相关的非满秩矩阵(测量矩阵)来对信号进行测量，该测量值的长度远小于原始信号的长度，即可利用测量值通过求解一个凸的最优化问题来实现原始信号的重构。

1.1 压缩感知理论

压缩感知过程实际上是用测量矩阵 A 对信号 x 进行“测量”，得到测量向量 y 。具体描述为：设 x 为离散信号， A 为 $m \times n$ 的测量矩阵，且 $m < n$ 。 $y = A \cdot x$ ，其中称 y 为测量信号。压缩感知过程如图 1。

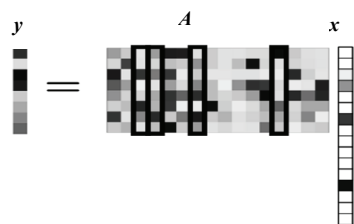


图 1 压缩感知过程

如果信号是可压缩的，那么当测量矩阵 A 满足 Johnson-Lindenstrauss^[5]准则时，能以最小误差为代价从 y 重构 x ，同时测量信号能保存原始信号的大部分信息。

1.2 特征提取和压缩

典型的测量矩阵是高斯随机测量矩阵，矩阵元素 $r_{ij} \in N(0 \sim 1)^{[6]}$ 。如果矩阵是非稀疏的且 m 比较大时，测量矩阵的存储和计算代价还是很大。笔者采用一个非常稀疏的测量矩阵，其元素定义如下

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

该矩阵中大部分元素都为 0，因此可以省去很多运算，能有效地降低算法计算复杂度。

跟踪特征质量直接影响算法性能。Haar 特征又称矩形特征，传统意义上的 Haar 特征模板内有白色和黑色 2 种矩形，Haar 特征值为白色矩形像素之和减去黑色矩形像素之和^[7]。

原始算法采用 Haar-like 特征的压缩域特征来

构建外表模型。Haar-like 特征区别于 Haar 特征，只利用矩形框内像素值总和作为特征。原始算法通过不同尺度的矩形框与样本进行卷积构建特征向量。由于特征向量维数很高；所以，利用前面提到的测量矩阵对特征向量进行降维。特征压缩过程如图 2。

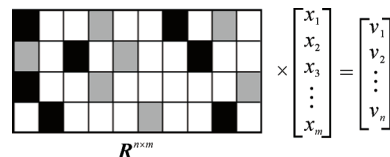


图 2 特征压缩过程

$R^{n \times m}$ 是测量矩阵， m 不同位置不同尺度的矩形框特征总数， $(x_1, x_2, x_3, \dots, x_m)^T$ 是原始高维特征向量，利用测量矩阵对其进行特征降维得到低维特征向量 $v = \{v_1, v_2, v_3, \dots, v_n\}^T$ 。

1.3 跟踪

跟踪过程被看成一个二元分类问题，利用正负样本的特征构造分类器，随后利用该分类器对候选样本进行判定，找到最优的样本作为最终跟踪结果。更新目标位置之后，同时采样提取新的目标和背景特征以更新分类器，为下一帧做准备。

2 算法改进

2.1 外表模型

外表模型的性能直接影响算法的性能。原始算法在整个样本区域随机提取一定数目的矩形框，这样的方式虽然能保证矩形框的随机性，但是这种随机方式形成的样本特征可能不能很好地表示样本特性而导致跟踪失败；因此，笔者提出了一种改进的特征构造方式。

对于输入样本，笔者将图像区域划分为 4 个互相之间有一定重叠的区域(灰色的矩形框)。将图像区域按这种方式划分，既能保证矩形框的随机性，同时矩形框又能分布在整个图像区域。避免矩形框随机性太强导致的特征表达性弱问题。

在每一个区域随机提取一个矩形框特征(黑色的小矩形框)，矩形框的位置和大小限定在所在的灰色区域内。特征的构造采取一种类似于对角 Haar 特征的计算方法，即将 2 和 3 灰色区域内的黑色矩形框像素之和减去 1 和 4 灰色区域内黑色矩形框像素之和作为特征向量的一维，即 $F = f_3 + f_2 - f_1 - f_4$ 。它的特征空间为 $\text{FeatureSpace} = \{F_1, F_2, F_3, \dots, F_m\}$ 。在这里 m 的数量级很大，特征向量维度很高，直接用

于跟踪会产生维度灾难。对于高维特征向量通过满足一定准则的稀疏矩阵对特征向量进行降维，降维后的特征向量保留大部分原始特征信息，这为特征压缩提供强大的理论支持。区域划分如图 3 所示。

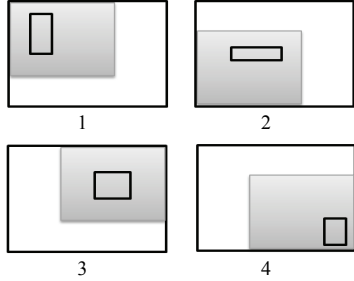


图 3 图像区域划分

不同颜色的矩形框表示对应矩形特征的系数，黑色矩形框表示特征系数为-1，灰色矩形框表示系数+1，白色矩形框表示系数为 0，特征提取过程如图 4 所示。

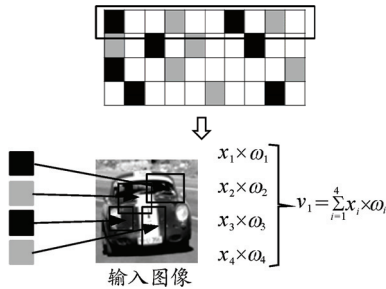


图 4 压缩特征获取过程

对于输入图像，从特征空间 FeatureSpace 内随机选择 s 个矩形特征。如图 4 中表格第一行共有 4 个非零元素(此时 $s=4$)且对应的特征值为 x_1, x_2, x_3, x_4 。特征对应的权重为 $\omega_1, \omega_2, \omega_3, \omega_4$ ，其中 $\omega_i \in \{-\sqrt{s}, 0, +\sqrt{s}\}, i=\{1,2,3,4\}$ 。外表模型的特征 F 是这 4 个 Haar-like 特征的线性加权组合^[8]，即 $v_1 = \sum_{i=1}^4 \omega_i f_i$ 。按照同样的方法得到 n 个特征，得到外表模型的特征向量 $\mathbf{v} = \{v_1, v_2, v_3, \dots, v_n\}^T$ 。这样就将原来的高维特征向量降维到低维向量，且降维后的特征向量保留原始特征向量的大部分信息。

2.2 分类器的构造和更新

分类器有很多种，如决策树、随机森林、支持向量机等。综合考虑计算成本、分类效果及学习算法实现难度。笔者采用基于样本各个特征相互独立假设的朴素贝叶斯分类器。对于任意样本 S ，它的压缩空间特征表示为

$$\mathbf{v} = \{F_1, F_2, F_3, \dots, F_n\}^T$$

通过贝叶斯分类器构建模型^[9]。

$$H(\mathbf{v}) = \log \left(\frac{\prod_{i=1}^n p(v_i | y=1)p(y=1)}{\prod_{i=1}^n p(v_i | y=0)p(y=0)} \right) = \sum_{i=1}^n \log \left(\frac{p(v_i | y=1)}{p(v_i | y=0)} \right) \quad (2)$$

在这里假定 $p(y=1)=p(y=0)$ ， y 表示样本的类别， $y=1$ 表示正样本， $y=0$ 表示负样本。Diaconis and Freedman 指出高维向量的随机投影向量总是符合高斯分布^[10]。因此分类器中的条件分布满足高斯分布，并有 4 个参数 $(\mu_i^1, \sigma_i^1, \mu_i^0, \sigma_i^0)$ ：

$$p(v_i | y=1) \sim N(\mu_i^1, \sigma_i^1) \quad (3)$$

$$p(v_i | y=0) \sim N(\mu_i^0, \sigma_i^0) \quad (4)$$

模型参数 $(\mu_i^1, \sigma_i^1, \mu_i^0, \sigma_i^0)$ 是增量更新的，更新方式为：

$$\mu_i^1 \leftarrow \lambda \mu_i^1 + (1 - \lambda) \mu^1 \quad (5)$$

$$\sigma_i^1 \leftarrow \sqrt{\lambda(\sigma_i^1)^2 + (1 - \lambda)(\sigma^1)^2 + \lambda(1 - \lambda)(\mu_i^1 - \mu^1)^2} \quad (6)$$

其中 λ 是学习参数，且 $0 < \lambda < 1$ 。

$$\mu^1 = \frac{1}{n} \sum_{k=0|y=1}^{n-1} v_i(k) \quad (7)$$

$$\sigma^1 = \sqrt{\frac{1}{n} \sum_{k=0|y=1}^{n-1} (v_i(k) - \mu^1)^2} \quad (8)$$

2.3 采样策略

正负样本的质量直接影响最终构造的分类器的判定能力。为了简单起见，笔者假定一个简单的运动模型，即目标相邻两帧的位置在一定范围内。通过 2 个距离阈值 r_{in} 和 r_{out} 界定采样的类别。候选样本和当前目标之间的距离小于阈值 r_{in} 是正采样，即 $SamplePos = \{Samples | distance < r_{in}\}$ 。若候选样本和当前目标之间的距离在 2 个阈值之间，则是负样本，即 $SampleNeg = \{Samples | r_{in} < distance < r_{out}\}$ 。 r_{in} 和 r_{out} 的大小控制正负样本的数目。

跟踪阶段大量候选样本都利用分类器进行判定，计算量比较大而影响算法实时性。为了提高算法速度，对候选样本进行预处理，将样本像素值方差作为一个预判定阈值进行预处理，即如果候选样本像素值方差低于目标像素值方差的一半，即认为该候选样本是目标样本的概率很小，可以排除^[11]。通过这样一个措施可以减少通过分类器进行判定的次数，从而减少大量运算，加快算法运行速度。

2.4 算法流程

初始化阶段 t -frame:

1) 设置初始参数: 生成稀疏矩阵、提取特征数目、分类器的初始参数等一系列设置。

2) 获取模板: 通过鼠标操作获得想要跟踪的兴趣区域, 即初始模板。

3) 提取特征: 对于正样本和负样本分别提取对应的压缩域特征。

4) 构造分类器: 利用得到的正负样本特征估计朴素贝叶斯分类器的相关参数, 构造朴素贝叶斯分类器。

跟踪阶段 $t+1$ -frame:

1) 获取候选样本: 基于上一帧目标的位置获取候选样本, 然后利用朴素贝叶斯分类器做二元分类处理, 得到最大分类器响应的样作为最终跟踪结果。

2) 更新分类器: 以更新过的目标位置为基准重新采样, 对新的正负样本提取特征, 重新估计分类器参数来更新分类器, 为下一帧做准备。

重复跟踪阶段的 1) 和 2) 步进行连续跟踪。

3 仿真结果及实验分析

实验仿真是基于 Visual Studio 2005 和 OpenCV

2.4 环境。

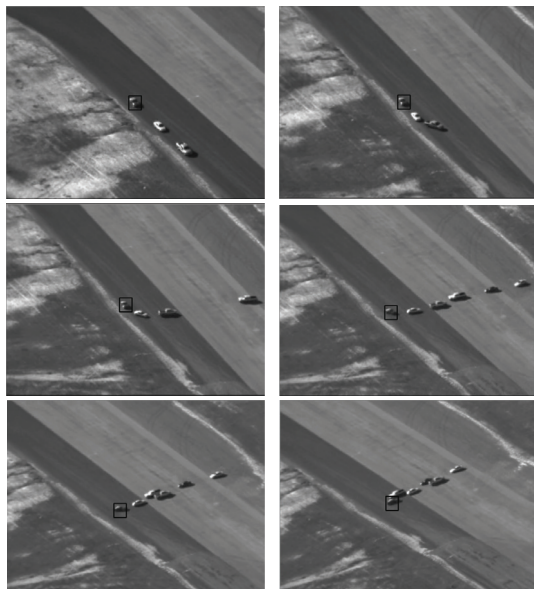


图5 空中对地车辆运动视频

如图 5 所示, 视频中的车辆开始自左上而右下开始运动, 中间改变方向自左下向右上运动。由于运动方向改变导致目标发生一定程度的旋转, 目标外表发生变化, 但是由于分类器是在线更新, 且模型更新充分利用背景信息, 因此跟踪器仍然能很好的跟踪目标。矩形框表示跟踪结果。仿真结果显示

该算法能处理一定程度的目标外表变化。

如图 6 所示, 视频中的灰色车自左向右运动, 途中经过车辆被绿色的树枝产生一定程度遮挡, 而跟踪器仍然能稳定准确的跟踪目标。仿真结果显示, 该算法对一定的遮挡具有一定的鲁棒性。

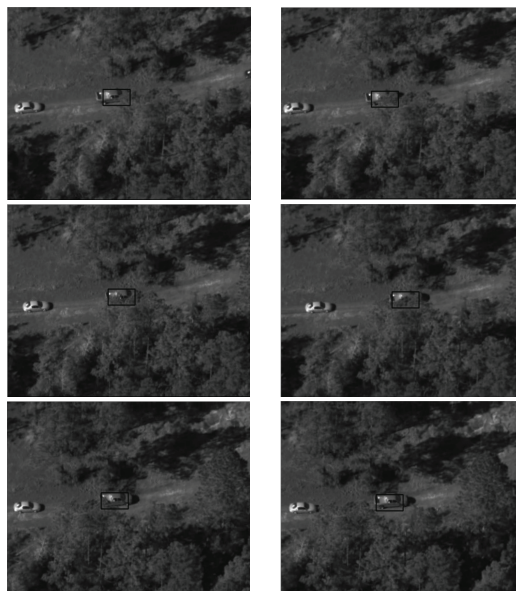


图6 有一定遮挡的空中对地车辆运动视频

算法可行性分析:

1) 更加稳定的图像特征: 采用 Haar-like 特征的加权线性组合作为图像特征, 该特征能处理一定的形变和光照变化。将 Haar-like 矩形框分散到整个样本区域, 避免原始算法由于随机取 Haar-like 特征而导致生成特征不能很好表示目标的情况。

2) 有效地特征降维: 通过非常稀疏的矩阵对原始特征向量进行降维, 既保持大部分原始目标的信息, 同时减少大量数学运算, 使算法在运行速度和跟踪效果之间达到很好的平衡。

3) 实时在线更新模型: 每次更新目标位置后都对正负样本重新采样, 通过新样本的特征重新估计分类器参数, 实时更新外表模型。通过在线更新的方式使分类器对目标和背景保持比较高的判别性, 因此该算法能处理一定程度的遮挡和旋转。

4 结论

笔者提出了一种改进的基于压缩特征的跟踪方法, 使用了一种新的特征构造方法; 同时, 引入基于稀疏矩阵的特征压缩方法降低特征向量的维数, 为算法的实时性提供可行性。仿真结果证明: 该算法能够有效地跟踪目标, 对目标外表变化和遮挡有一定适应能力, 且算法达到了很好的实时性。