上海大学无人艇工程研究院

——环境感知组

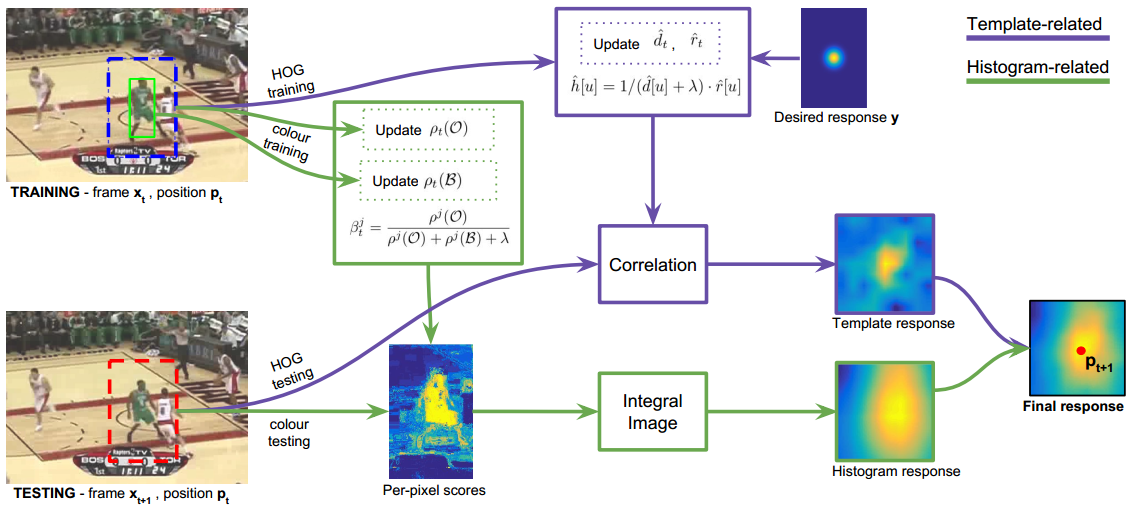
# Staple: Complementary Learners for Real-Time Tracking—Staple

作者：**Luca Bertinetto** Jack Valmadre Stuart Golodetz Ondrej Miksik Philip H.S. Torr

主页：<http://www.robots.ox.ac.uk/~luca/staple.html>

出处：2016年CVPR

源码：matlab



注：**加粗**的作者为重点关注研究者

图注：本算法的核心示意图

Date：2017.11.30

## 版本更新记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日 期** | **更新人** | **主要更新内容描述** | **版本号** |
| 2017年03月28日 | 陈加宏 | 完成大致框架搭建 | V1.0.0 |
| 2017年09月26日 | 陈加宏 | 完成算法细节的总结 | V1.0.1 |
| 2017年11月30日 | 陈加宏 | 完成算法理解的细化 | V1.0.2 |

目 录

[Staple: Complementary Learners for Real-Time Tracking—Staple 1](#_Toc499795668)

[版本更新记录 2](#_Toc499795669)

[1、概述 3](#_Toc499795670)

[1.1 前言 3](#_Toc499795671)

[1.2 创新点 3](#_Toc499795672)

[2、细节 4](#_Toc499795673)

[2.1 主要流程 4](#_Toc499795674)

[2.2 数学模型 6](#_Toc499795675)

[2.3 模型求解 6](#_Toc499795676)

[3、实验 6](#_Toc499795677)

[3.1 代码框架 6](#_Toc499795678)

[3.2 实验结果及分析 6](#_Toc499795679)

[3.3 优缺点总结 6](#_Toc499795680)

[3.4 今后工作 6](#_Toc499795681)

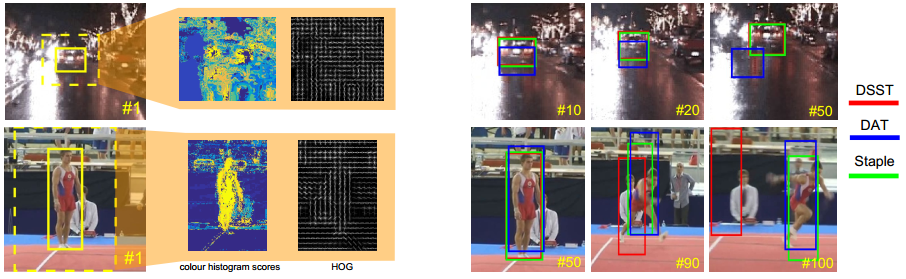
## 1、概述

该部分主要讲述的是颜色特征对于视觉目标跟踪算法的性能增强。特征适用于表示目标的定量数据，越好的特征越能用来表示目标。在目标跟踪领域意味着使用该特征来表示目标更加具有可分性，可以使用特征来判断目标还是背景，目标还是相似目标，简单的说，越好的目标越有利于跟踪算法的运行。已经有研究表明，特征表示是目标跟踪各个环节中最重要的环节，没有好的特征表示，该算法肯定不会很好，所以这部分论文对于目标特征的增强对于整个跟踪算法性能的提升具有显著意义。

### 1.1 前言

在2010年MOSSE横空出世之后，相关滤波技术在目标跟踪领域可算是刮起了一阵旋风，目标的特征表示也从最开始的单通道原始灰度特征扩展到多通道的HOG特征，HOG属于模板特征，在2005年提出用于行人检测，主要能够检测出目标的轮廓，对于光照信息不敏感，在目标跟踪算法中表现优异。在这之后，使用的增强特征更多的关注颜色方面，这也比较容易理解，颜色特征在视觉感知中占有十分重要的地位，人类的视觉感知有很大一部分借助于颜色。接下来介绍颜色特征。

颜色特征，在目标跟踪中颜色是个非常重要的特征，不管多少个人在一起，只要目标穿不用颜色的一幅就非常明显。前面介绍过2014年CVPR的CN是相关滤波框架下的模板颜色方法，还有基于的统计颜色特征方法DAT，可能是由于相关滤波中HOG的效果太好导致在传统的特征设计方面能够使用的特征并不多，而颜色特征是HOG的有效补充。



前面分析了相关滤波模板类特征HOG对快速变形和快速运动效果不好，但对运动模糊光照变化等情况比较好；而颜色统计特征颜色直方图对变形不敏感，而且不属于相关滤波框架没有边界效应，快速运动当然也是没问题的，但对光照变化和背景相似颜色不好。综上所述，这两类方法可以互补，也就是说DSST和DAT可以互补结合。上图是原作者结合DSST和DAT得到的Staple算法同时具备了两种算法的优点，对于容易受背景颜色信息干扰的场景，Staple可以发挥DSST的强项克服光照颜色变化的影响，对于目标的快速变形，该算法可以发挥DAT对于快速变形不敏感的优势，并且速度还比各个子算法快，所以这篇算法是相当好的实时性跟踪算法的baseline。Staple算法的全称为Sum of Template And Pixel-wise Learners跟踪算法。

### 1.2 创新点

1. 考虑了两种方法的结合。Learning the template score（Hog特征）+ Learning the histogram score（简单的颜色直方图）

2. 相关滤波用**HOG特征时对运动模糊和照度很鲁棒，但是对形变不够鲁棒**。而**颜色直方图对形变则非常鲁棒**。毕竟一个目标有了形变后，整个目标的颜色分布是基本不会变的。另一方面，颜色直方图对光照变化不鲁棒，这一点又可以由HOG特征进行互补。

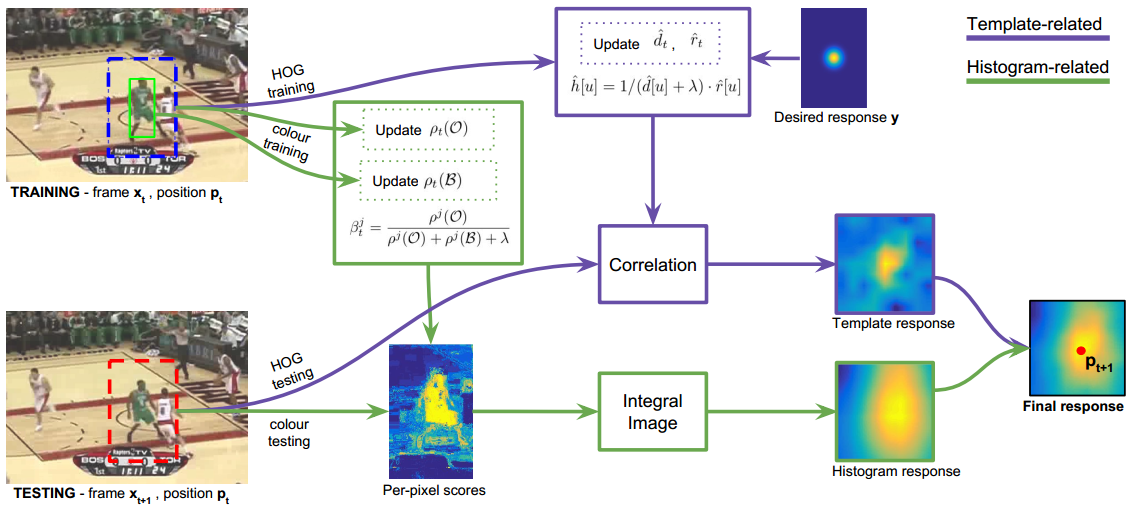
3. **整体新意就在于两种方法的结合**，互补之后效果鲁棒，速度很快，因为相关滤波用在跟踪中，可以用FFT快速傅里叶变换进行加速，如KCF的速度可以到140FPS，这篇文章中提到的他的速度为**80FPS**。

4. **效果在OTB50上还不错，在VOT14和VOT15上非常好**。作者给出的OTB50的结果在定位精度（precision）这个指标不是那么靠前，所以在文章中并没给出，但是overlap这个指标还不错，这是因为除了算法本身对结果的提升，作者还有考虑**对尺度的处理，**用的是尺度金字塔方法，只加到HOG上了，颜色特征没有加尺度，因为对颜色加尺度意义不大，反而会让结果更差。

## 2、细节

如果要用一句话介绍Luca Bertinetto(牛津大学)的Staple tracker，那就是把模板特征方法DSST(基于DCF)和统计特征方法DAT结合。该部分主要讲述本文算法的核心细节，包括算法的主要流程、数学模型的建立以及模型的求解方法。要完全的理解跟踪算法必须从最基础的问题本质出发，借助数学模型对问题进行抽象，最后通过优化求解方法得到解决方案。

### 2.1 主要流程



这篇文章整体思路比较简单，是将DSST那一套加了颜色直方图。考虑到基于HOG特征的相关滤波跟踪器（比如KCF，DSST）对**运动模糊和光照变化鲁棒性高**，而对形变的鲁棒性较低，在目标发生比较严重的形变时容易丢失目标，如上图所示，图中的DSST在遇到形变大的时候，容易丢失目标。总的来说，本文是**基于DSST并且加入对形变鲁棒的颜色直方图特征滤波器。**

简单吗？非常不简单，两个框架的算法高效无缝结合，将25FPS的DSST和15FPS的DAT结合后速度竟然达到了80FPS。DSST框架把跟踪划分为两个问题，即平移检测和尺度检测，DAT就加在平移检测部分，相关滤波有一个响应图，像素级前景概率也有一个响应图，两个响应图线性加权得到最终响应图，其他部分与DSST类似，平移滤波器、尺度滤波器和颜色概率模型都以固定学习率线性加权更新。要理解该算法需要解决以下几个问题：

1、既然是颜色特征，DSST+DAT换成DSST用HOG+CN一样吗？

肯定是不一样的。1. CN特征是模板类特征，DAT中统计颜色直方图是概率特征，举个例子，跟踪目标是人，穿白色上衣，黑色裤子，CN特征学习到上白下黑的模板，颜色直方图学习到目标60%是白色，40%是黑色。如果这个人倒立了，CN模板对不上就认为这个不是目标，而颜色直方图依然60%白40%黑，判定就是目标。2. HOG+CN任然在相关滤波框架中，具有相关滤波的固有缺陷，而像素级概率颜色概率响应不受这一影响，能正确检测快速变形和快速运动情况，**某种程度上来说DAT缓解了CF的边界效应**，所以对快速运动情况不像DSST那么差。

2、Staple为什么能控制计算量，比DSST和DAT都快还不损失性能？

首先，Staple中检测和更新的图像块大小限制在150\*150以内，即HOG特征分辨率在38\*38以内，这样DSST的速度就会很快。但前面也分析过即使2因子下采样精度已经比较低了，那为什么Staple的精度不受影响呢？因为平移相关滤波响应和颜色概率响应要线性加权，而**颜色概率响应得分辨率在150\*150左右下降不是很多**，这就弥补了平移相关滤波的精度损失。举个例子，原始检测图像块分辨率是400\*400，如果限制分辨率到100\*100，HOG特征分辨率是25\*25，响应图也是25\*25，1像素位移对应跟踪框16像素位移；而加入DAT后，颜色概率响应分辨率是100\*100，相关滤波响应先上采样到100\*100，再两个响应结合，最终响应的分辨率是100\*100，1像素位移对应跟踪框4像素位移，跟踪精度提高了很多且保持低计算量。其次，**像素级颜色概率图到颜色概率响应采用积分图加速计算**，积分图是加速神器不用多说。这两种策略分别提高了DSST和DAT的速度，而且Staple并没有破坏封闭解，所以非常非常之NIUBILITY。

3、两种响应图线性加权，会不会太简单，效果不够好？

具体实现中采用比较简单的线性加权，最终响应 = 0.7\*相关滤波响应 + 0.3\*颜色概率响应，可以看出Staple更加相信相关滤波的结果，而颜色概率只是起到辅助作用，看代码作者实现了另一种加权方式，可能效果不好最终没用。

相关滤波响应值是判别器输出的置信度，可以表示平移滤波器的可靠程度，跟踪越可靠峰值越大；颜色概率响应值**表示当前像素点属于前景的置信度**，像素点及附近像素属于前景的概率越高，响应值越大，也可以表示颜色概率模型的可靠程度。所以作者说这种线性加权是置信度加权，CF代表相关滤波的峰值点，PWP代表颜色概率响应的峰值点：

1、如果CF和PWP大小相似时，以相关滤波响应为主，颜色响应仅微调跟踪框的位置；

2、如果CF比较大，PWP比较小时，说明相关滤波响应更可靠，此时颜色概率响应几乎没有贡献；

3、如果CF比较小，PWP比较大时，说明颜色响应更可靠，以颜色峰值所在区域的CF局部峰值为主；

以上就是非常推荐而且比较快的Staple，在VOT数据库上性能非常高，竞赛结果已经看到了，前面也提到过颜色特征方法在OTB上不会很好，在OTB50上Staple略低于SRDCF算法。Staple虽然简单高效，但也有不足之处，例如，仅以峰值判断置信度并不可靠，以上策略并不总是正确的。如果能有更全面可靠的指标，让我们知道相关滤波的检测置信度，那**加权系数就可以自适应确定**，在相关滤波检测置信度较低时更多地相信颜色概率的结果，也许会更加合理。也可以用更好的模板特征和统计特征在这个框架中完美结合，提升性能。

### 2.2 数学模型

### 2.3 模型求解

## 3、实验

该部分主要讲述算法实现代码的主要流程、实验环境及效果分析、算法优缺点的总结，最后提出后续可改进的方面。实验是检验真理的唯一标准，那么对实验结果详细的分析以及结合算法的原理对算法本质上的一些思考有利于之后研究工作的开展，也是今后工作的一个研究突破点。

### 3.1 代码框架

### 3.2 实验结果及分析

### 3.3 优缺点总结

### 3.4 今后工作