上海大学无人艇工程研究院

——环境感知组

# Context-Aware Correlation Filter Tracking

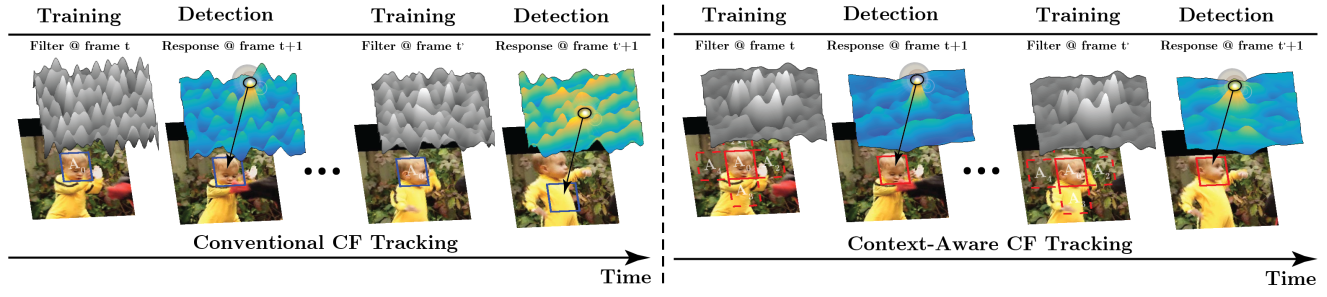
# CFCA

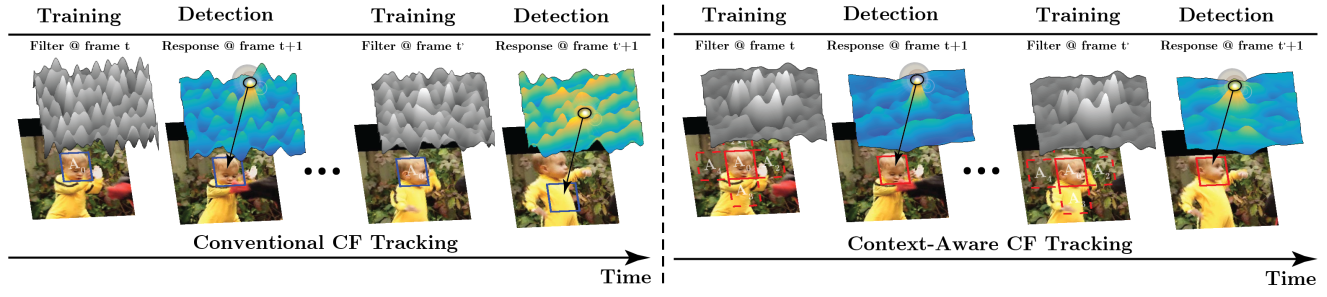
作者：**Matthias Mueller** Neil Smith Bernard Ghanem.

主页：<https://ivul.kaust.edu.sa/Pages/pub-ca-cf-tracking.aspx>

出处：2017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：2019.02.20

## 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 |
| 2017年09月22日 | 陈加宏 | 完成算法细节的总结 | V1.1 |
| 2017年11月21日 | 陈加宏 | 完成算法的理论推导 | V1.2 |
| 2019年02月20日 | 陈加宏 | 修改 | V2.0 |

目 录

[Context-Aware Correlation Filter Tracking 1](#_Toc1584626)

[CFCA 1](#_Toc1584627)

[版本更新记录 2](#_Toc1584628)

[1、概述 3](#_Toc1584629)

[1.1 前言 3](#_Toc1584630)

[1.2 创新点 5](#_Toc1584631)

[2、细节 5](#_Toc1584632)

[2.1 主要流程 5](#_Toc1584633)

[2.2 数学模型 5](#_Toc1584634)

[2.3 模型求解 6](#_Toc1584635)

[3、实验 8](#_Toc1584636)

[3.1 实验结果及分析 8](#_Toc1584637)

[3.2 优缺点总结 8](#_Toc1584638)

[3.3 今后工作 8](#_Toc1584639)

## 1、概述

该部分主要讲述的是视觉目标跟踪的研究背景以及相关滤波CF算法提出的意义，同时分析相关滤波算法的优缺点以及针对CF类算法的缺点提出可选择的改进方法。结合经典CF类算法存在的部分问题讲述本文算法的创新点以及具体解决方法的概述。

### 1.1 前言

相关滤波CF类跟踪算法起源于2010年的开创性工作MOSSE，该算法第一次将相关滤波技术从信号处理迁移至视觉跟踪领域，其最核心的创新点在于将目标图像与滤波器的相关操作通过傅里叶变化后转移到频域中进行求解，将卷积操作转换成了按元素点乘，这样算法的计算复杂度大大的降低了，从而实现了该算法速度极快（669FPS）的效果。但是MOSSE算法的样本采样仍是一种稀疏采样，训练效果一般；同时采用的分类器是线性滤波器（最小二乘法），分类性能一般；采用的特征是单通道的灰度特征，表征目标的能力有限。针对上述的缺点，2012出现的CSK以及2014年的KCF对其进行了针对性的改进，采用的分类模型是具有更好分类属性的岭回归（在最小二次的基础上加入正则化项来避免模型训练的过拟合），同时从岭回归的闭式解出发，为了避免大规模矩阵的求逆操作开创性的利用了循环矩阵的特性，提出训练样本的循环位移的假设来近似一种密集的训练样本采样，这样不仅能够保持很好的训练速度，同时训练效果大幅度的提高了；原文作者为了将多维特征（如HOG）以及扩展线性滤波器来进一步提升模型的分类性能，把岭回归在基本空间的求解利用转换到对偶空间，进一步证明核函数也满足循环矩阵的性质，那么直接根据循环矩阵在频域中的对角化来实现对算法求解的简化。以上就是CF类算法的baseline，从此确立了CF的实现框架，由于CF算法在实时性方面杰出的表现，一举奠定了其在视觉目标跟踪的历史地位，并在2010年开始至今（2017）仍保持很高的研究热度，基于KCF的框架来实现工业视觉跟踪的技术已经相当的成熟。CF的baseline框架直接解决了视觉跟踪的实时性要求，这样研究人员就会在其基础上有针对性的改进其跟踪精度和鲁棒性。

总体来说，相关滤波类方法对**快速变形和快速运动**情况（还有背景中存在类内相似物时效果不好）的跟踪效果不好。

快速变形主要因为CF是模板类方法，容易跟丢。这个比较好理解，前面分析了相关滤波是模板类方法，如果目标快速变形，那基于HOG的梯度模板肯定就跟不上了，如果快速变色，那基于CN的颜色模板肯定也就跟不上了。这个和模型**更新策略与更新速度**有关，固定学习率的线性加权更新，如果学习率太大，部分或短暂遮挡和任何检测不准确，模型就会学习到背景信息，积累到一定程度模型跟着背景私奔了，一去不复返。如果学习率太小，目标已经变形了而模板还是那个模板，就会不认识目标。举个例子，比如多年不见的同学，你很可能就认不出了，而经常见面的同学，即使变化很大你也认识，因为常见的同学在你大脑里面的模型在持续更新，而多年不见就是很久不更新。

快速运动主要是边界效应(Boundary Effets)，而且边界效应产生的错误样本会造成分类器判别力不够强，下面分训练阶段和检测阶段分别讨论。训练阶段，合成样本降低了判别能力。如果不加余弦窗，那么移位样本是长这样的：

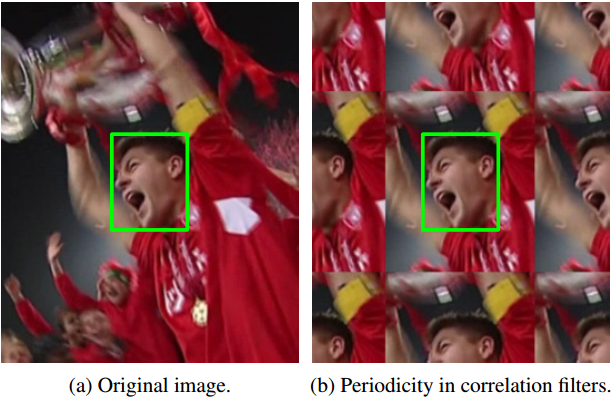


图1 核相关滤波算法中循环假设的示意图

除了那个最原始样本，其他样本都是“合成”的，100\*100的图像块，只有1/10000的样本是真实的，这样的样本集根本不能拿来训练。如果加了余弦窗，由于图像边缘像素值都是0，循环移位过程中只要目标保持完整，那这个样本就是合理的，**只有目标中心接近边缘时，目标跨越边界的那些样本是错误的**，这样虽不真实但合理的样本数量增加到了大约2/3(padding= 1)，即使这样仍然有1/3(3000/10000)的样本是不合理的，这些样本会降低分类器的判别能力。同时加余弦窗并不是免费的，余弦窗将图像块的边缘区域像素全部都变成0，大量过滤掉分类器本来非常需要学习的背景信息，原本训练时判别器能看到的背景信息就非常有限，我们还加了个余弦窗挡住了背景，这样进一步降低了分类器的判别力。

检测阶段，相关滤波对快速运动的目标检测比较乏力。相关滤波训练的图像块和检测的图像块大小必须是一样的，这就是说你训练了一个100\*100的滤波器，那你也只能检测100\*100的区域，如果打算通过加更大的padding来扩展检测区域，那样除了扩展了复杂度，并不会有什么好处。目标运动可能是目标自身移动，或摄像机移动，按照目标在检测区域的位置分四种情况来看：

如果目标在中心附近，检测准确且成功。

如果目标移动到了边界附近但还没有出边界，加了余弦窗以后，部分目标像素会被过滤掉，这时候就没法保证这里的响应是全局最大的，而且，这时候的检测样本和训练过程中的那些不合理样本很像，所以很可能会失败。

如果目标的一部分已经移出了这个区域，而我们还要加余弦窗，很可能就过滤掉了仅存的目标像素，检测失败。

如果整个目标已经位移出了这个区域，那肯定就检测失败了。

以上就是边界效应(Boundary Effets)，而本文试图提升CF类框架中的跟踪鲁棒性，主要从循环假设带来的边界效应角度出发。再次强调这些方法速度相对比较慢，相关滤波傲视群雄的高速已经不见了，但换来了可以匹敌深度学习方法的性能。

### 1.2 创新点

研究动机：近年来，由于相关滤波能够在保证较高精度的情况下保证速度，基于CF的目标跟踪算法层出不穷。对CF在跟踪中的应用不了解的可以去看KCF。这些算法都基于两个前提：1、背景是单一的或者说均匀的；2、经过循环移动的样本跟真实的样本相等。

一般第一点通过加cosine窗实现，同时也能抑制边缘效应，第二点其实做不到，只能近似相等了。为了不让跟踪结果出现漂移的意外情况，一般的算法的搜索范围在上一帧目标的2倍左右的大小，能利用的背景信息很少，另外在加了cosine窗之后，背景的信息再次被减少了。因此这里就是作者要解决的问题了，基于CF的跟踪算法通常只有有限的背景信息，对快速运动，遮挡等的跟踪效果会受限制，如何能够合理地增加背景信息，又不带来负面的影响呢？

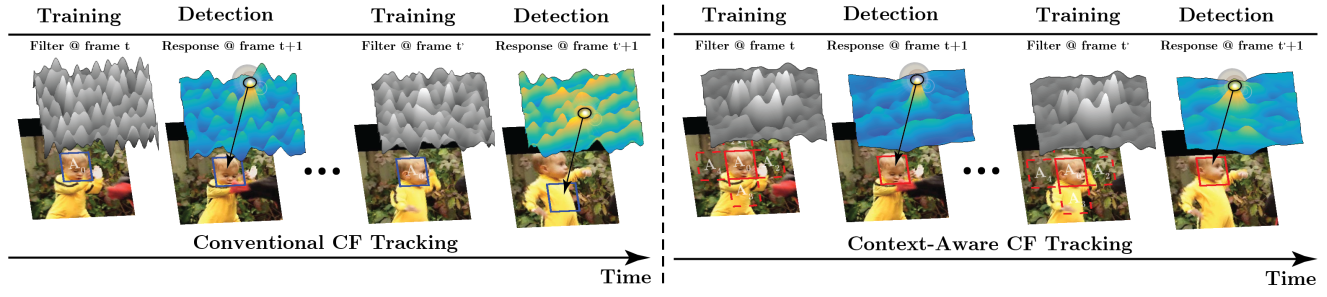


图2 CFCA算法与传统CF类算法的原理对比示意图

解决方法：作者提出了一种新的框架，能够增加更多背景的信息，并把这些global context包含到学到的filters里面。经过推导，提出的这个框架也是有闭环解的，并且速度的影响相对较小。如上图所示，左边是传统的基于CF的算法，右边是作者提出的context-aware CF（本文简称为CA-CF）的算法。可以看到增加了上下左右四块context之后，在目标有了明显的旋转变化后，依旧跟的很好，从他们的response map来看，加context之后的相应图的干扰也有较少，判别力更强。

## 2、细节

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

本文提出的CFCA算法属于框架类的CF跟踪方法，原作者提出的框架可以应用于所有CF类算法，作者将CA框架应用于MOSSE、DCF、SAMF以及Staple四个相关滤波跟踪算法上，用以验证本文提出的上下文感知跟踪框架的有效性。

### 2.2 数学模型

为了研究的完整性，先对传统的CF框架做出求解的推导，根据前面的介绍， 已经很熟悉这样的一个公式，岭回归：

(1)

对上式的求解可以直接在空域中也可以在频域中进行，可以在基本空间中求解，也可以在对偶空间中求解。一般来说为了可以加入对多维特征和非线性分类的支持会将求解过程放在频域对偶空间中求解，这样可以利用循环矩阵在频域中可以对角化的性质来简化求解过程。而本文考虑到在模型训练中加入更多的背景信息，来提高分类器的判别能力，所以在原来的岭回归基础上加入了目标空间上下文的约束，这样优化目标函数就变成了：

(2)

原文作者从单通道特征和多通道特征两个方面对其进行推导求解，可以得到闭式解。

引入更多的背景信息进行相关滤波器的学习，的确对效果有改进。关键的是如何从原理上把背景信息利用起来，这篇文章的做法挺巧妙，实验证明了其效果。下一小结分别从单通道和多通道两个方面来说明对本文优化目标函数的求解。

### 2.3 模型求解

开始还是以传统的CF类框架的求解推导过程来说明岭回归从基本空间和对偶空间的求解过程。对1式求导后置0可得：

（3）

为了避免对高维矩阵进行计算机极大的求逆操作，因为是循环矩阵，那么可以对进行频域对角化处理：

对上式进行频域变换得到：

在目标检测时，跟踪已经训练得到的滤波器参数和新到来的带检测样本，可以根据下式获得响应值：

在对偶空间中，可以将表示为，那么对于的求解相当于对的求解，将式3推导成这样：

继续使用频域对角化循环矩阵，可以得到：

检测过程如下：

接下来进入本文算法的核心求解部分，也就是对加入了新约束的岭回归公式2进行基本空间和对偶空间的求解过程。上式的求解分为单通道特征和多通道特征的求解，下面分开来叙述：

单通道特征的情况：

上式可以写成：

那么对上式的求解还是一个凸优化问题，可以求出闭式解：

还是同样的套路，应用循环矩阵对角化可以避免矩阵求逆：

以上是基本空间的求解，在对偶空间中：

继续使用循环矩阵的对角化性质，可以得到：

检测过程如下：

多通道特征的情况：

对上述优化函数重写成如下形式：

上式同样是凸优化问题，可以求出闭式解：

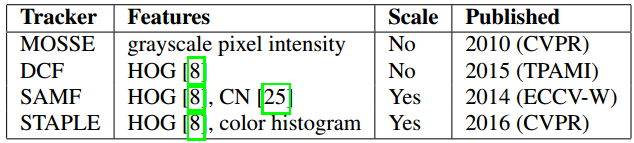
所有的求解过程与单通道类似，具体的代数代入会不所不同，这里不详细叙述。

## 3、实验

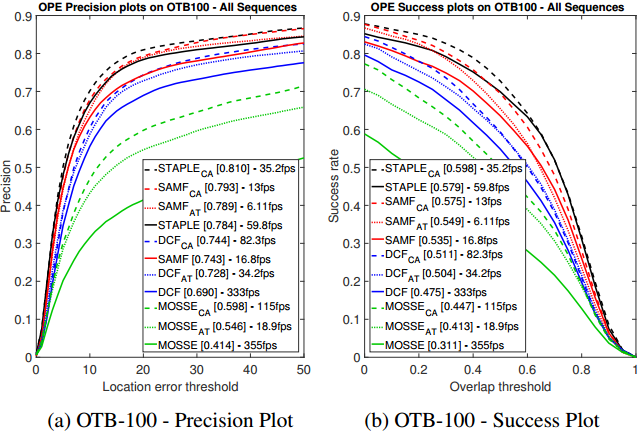
该部分主要讲述算法实现代码的主要流程、实验环境及效果分析、算法优缺点的总结，最后提出后续可改进的方面。实验是检验真理的唯一标准，那么对实验结果详细的分析以及结合算法的原理对算法本质上的一些思考有利于之后研究工作的开展，也是今后工作的一个研究突破点。

### 3.1 实验结果及分析

用来做对比实验的CF类跟踪算法有4个：



跟踪性能的评估还是依据OTB的评估体系。框架改变后的实验对比结果如下所示：



### 3.2 优缺点总结

这篇CVPR2017最新的文章，很明显，这是一篇从框架角度来提升CF类跟踪体系的跟踪鲁棒性的，在目标函数中加入了更多的背景干扰信息，有效的提升了跟踪的鲁棒性，本文最重要的贡献是在加入了更多背景信息的前提下，推导出目标函数的闭式解，使得整个框架的执行效率受到的影响不大，可以保证一定的跟踪实时性。

### 3.3 今后工作

本文关于增多的上下本信息的选取策略是我关注的方向，从代码来看，本文选用的上下文信息是与目标框大小一致的，位于目标框前后左右的4个背景信息，原作者在这部分的讨论中提到：这种上下左右的采样策略可以平均的提升目标检测的鲁棒性，还可以在上下文信息的选取策略上选取卡尔曼滤波估计的目标运行来采样，也可以结合2015ICCV上的文章SOWP来做针对性的改变，这是一个可以出论文的思路，可以马上做实验尝试。