上海大学无人艇工程研究院

——环境感知组

# High-Speed Tracking with Kernelized Correlation Filter

# KCF

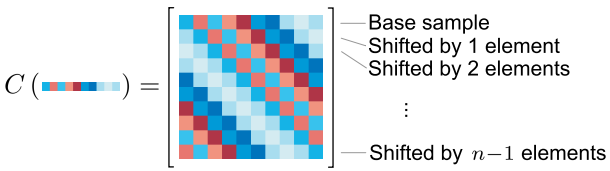
作者：**João F. Henriques** Rui Caseiro Pedro Martins Jorge Batista

主页：<http://www.robots.ox.ac.uk/~joao/#publications>

出处：2014年PAMI

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：2019.02.15

## 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25日 | 陈加宏 | 完成算法细节的总结 | V1.1 |
| 2017年11月21日 | 陈加宏 | 完成算法细节的修改 | V1.2 |
| 2019年02月15日 | 陈加宏 | 修改框架 |  |

目 录

[High-Speed Tracking with Kernelized Correlation Filter 1](#_Toc1112252)

[KCF 1](#_Toc1112253)

[版本更新记录 2](#_Toc1112254)

[1、概述 3](#_Toc1112255)

[1.1 前言 3](#_Toc1112256)

[1.2 创新点 4](#_Toc1112257)

[2、细节 4](#_Toc1112258)

[2.1 主要流程 4](#_Toc1112259)

[2.2 数学模型 5](#_Toc1112260)

[2.3 模型求解 8](#_Toc1112261)

[3、实验 11](#_Toc1112262)

[3.1 代码框架 11](#_Toc1112263)

[3.2 实验结果及分析 11](#_Toc1112264)

[3.3 优缺点总结 11](#_Toc1112265)

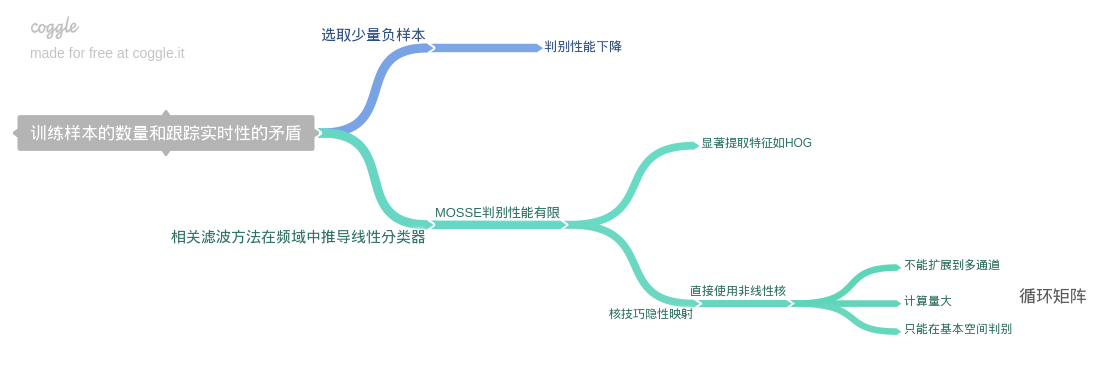
[3.4 今后工作 11](#_Toc1112266)

## 1、概述

KCF(Kernelized Correlation Filters)称为核化相关滤波器，原文作者在2012年ECCV会议上首次提出CSK，2014年，作者从多通道HOG特征、实验结果对比分析和解释，公式推导等几个方面对CSK进行扩展描述，将其发表于TPAMI上。作者利用Yi Wu等人提供的基准跟踪测试视频OTB，同Struck和TLD等当时优秀主流跟踪方法从跟踪性能和处理速度进行对比。KCF方法在实现更好跟踪性能的条件下，处理速度可达300fps，约为一般跟踪算法十倍，并且实现简便。

## 2、问题分析

大多现代跟踪器的核心部分为判别式分类器，以对目标和其所处环境进行判别。为提升判别式分类器性能，以应对视觉目标跟踪中所面对的各种挑战，满足跟踪性能需求，需要利用大量各式样本，如平移和缩放样本，对其进行在线训练和判别。在实际应用过程中，大多跟踪算法属于在线跟踪算法需满足实时要求。基于判别式分类器目标跟踪算法需迭代利用跟踪结果选取学习正负样本。学习样本过多制约实时性，学习样本过少降低跟踪性能。为此基于判别式分类器跟踪算法**实时性和样本数量决定的算法性能之间存在很深矛盾**。为解决此矛盾，通常有两种方式：一种选取少量负样本以降低跟踪算法性能；另一种是采用相关滤波器方法即利用时域卷积等于频域点乘属性，从频域推导用于样本平移判别的线性分类器。对于第一种方法，Joao F. Henriques认为，影响基于判别式分类器跟踪算法性能主要因素是选取的样本数量太少，影响了分类器的性能。

****

1. MOSSE算法到KCF改进的思路过程示意图

虽然相关滤波器MOSSE实时性和学习样本个数选取之间矛盾较小，但其是在基本空间实现判别，判别性能有限。为解决线性相关滤波器分类性能不足问题，有两种途径：通过显性提取待跟踪目标特征HOG，基于目标特征对目标进行跟踪；通过核方法隐性将原始像素映射至高维特征空间。对于显性特征提取方法其研究重点转换为另外一个研究方向即特征提取。另外，线**性相关滤波器只能应用于单通道特征目标跟踪**，其还没有扩展至多通道或者扩展至多通道无法满足实时跟踪要求。对于核方法，虽然已有人提出了Kernel Synthetic Discriminant Function(KSDF)用于目标判别，但是没有建立平移样本、非线性核和傅里叶变换之间关系。若将非线性核方法拓展至相关滤波器会导致：计算量十分巨大；只能在基本空间而无法在对偶空间进行判别；无法扩展至多通道。为此作者通过循环矩阵，利用傅里叶变换建立平移样本和快速核学习检测方法之间的联系，解决由于样本个数缺少导致的学习算法性能和实时性之间的矛盾，从而实现高速目标平移位置跟踪。

**KCF将训练样本和参考样本表征为循环矩阵形式，利用循环矩阵块性质对其进行对角化，将完备平移样本在核空间即高维空间相关匹配即卷积问题转换为一个基准样本间在频域点乘问题，而无需遍历所有平移样本进行相关匹配，对当前样本块和原先样本块在高维特征空间进行快速匹配，从而实现目标位置跟踪。**

## 3、解决方法

循环矩阵可以通过基准样本进行循环平移进行构建，表征了样本的平移，并且能通过统一的离散傅里叶变换正交基对其进行对角化。循环样本可以被基准样本表征，**完备平移样本即基准样本平移生成的所有样本在空域相关即卷积计算，可以转换为基准样本在频域的点乘计算**。KCF仅消耗同线性相关滤波器相当的计算量就可以解决非线性核相关滤波器扩展问题，极大降低平移样本间相关性计算所需计算量和存储量。通过观测样本和匹配样本所对应所有平移样本之间的相关计算，选取最大相关性所对应位置以跟踪目标位置变化。

为快速隐性实现样本在核空间的相关计算，需构建循环核矩阵(块)。以基准样本x构建循环样本矩阵。若某一核函数k为**酋不变核函数**，则以其为映射关系构建高维空间对应的核矩阵K也为循环矩阵，称为**循环核矩阵**。当核矩阵为循环矩阵时，能实现快速核回归(参数学习)、快速检测(相关匹配)。快速核回归和快速检测计算都分别需要计算核相关量k。核相关量k为循环核矩阵的基准量。当核函数选为**点积核函数或径向基核函数**时，能通过快速傅里叶变换实现的快速计算，否则需要**通过滑窗平移**计算核相关量。

KCF实现主要分为四步，分别为初始化；模型参数学习(核回归)；检测；参数和参考样本更新。其中核回归和检测属于核心部分。下面给出四个步骤的详细描述。

1、初始化：初始化包括跟踪目标位置，大小初始化，算法参数初始化，观测样本和参考样本初始化。观测样本和参考样本的选取方法为：以待跟踪目标为中心，选取比目标尺寸大若干倍的区域分别作为匹配样本和观测样本的基准样本z和x。

2、核回归。核回归功能根据标记样本实现判别式分类器参数学习。以上一时刻匹配位置为中心获取k时刻观测样本的基准样本及其回归值为输入，利用酋不变核函数构建匹配好观测样本的核自相关量，实现对分类器参数学习。

3、检测。读取图像数据获得k+1时刻观测样本x，利用酋不变核函数构建匹配样本和观测样本核相关量，进而进行平移相关匹配，得到相关量最大平移量即跟踪目标状态。

4、模型参数和参考样本更新。利用匹配好的观测样本和学习回归参数分别同已有匹配样本和回归参数进行加权获取最新的匹配样本和回归参数。k=k+1，返回第2步。

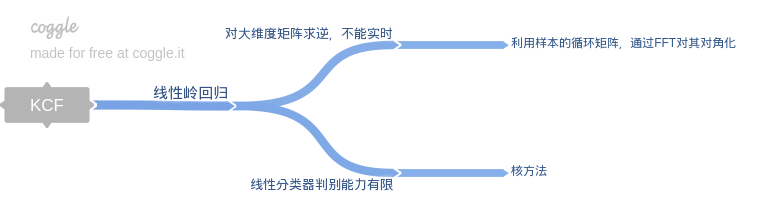


图2 KCF原理分析图

**创新点：**

KCF是一种判别式追踪方法，这类方法是在追踪过程中训练一个目标检测器，使用目标检测器去检测下一帧预测位置是否是目标，然后再使用新检测结果去更新训练集进而更新目标检测器。KCF的主要贡献：

1、使用目标周围区域的循环矩阵采集正负样本，利用脊回归训练目标检测器，并成功的利用循环矩阵在傅里叶空间可对角化的性质将矩阵的运算转化为向量的Hadamad积，即元素的点乘，大大降低了运算量，提高了运算速度，使算法满足实时性要求；

2、将线性空间的脊回归通过核函数映射到非线性空间，在非线性空间通过求解一个对偶问题和某些常见的约束，同样的可以使用循环矩阵傅里叶空间对角化简化计算；

3、给出了一种将多通道数据融入该算法的途径。

## 4、原理分析

**数学模型：**

基于判别式分类器目标跟踪问题可以表述为回归问题。首先分别介绍线性岭回归模型和基于核函数的非线性岭回归模型及其特性。然后分析利用回归模型实现快速目标跟踪时，矩阵样本和核矩阵需满足一定特性，从而引入**循环样本矩阵**和**循环核矩阵**概念，推导出核相关滤波跟踪器并通过DFT给出快速参数学习，快速检测过程的**闭合求解公式**。参数学习和检测都需求解对应的**核相关量**，当核函数为点积核函数和径向基函数时，通过DFT核相关量也存在快速闭合求解公式。

**线性岭回归：**

线性回归的目的是通过样本集，建立变量X和y之间线性映射关系。线性岭回归建立如下目标函数并关于参数w进行极小化求解。

以上公式第一项是一个均方差项，第二项是一个规则化项（也叫**权重衰减项**），其目的是减小权重的幅度，防止过度拟合。λ控制规则化权重大小。线性岭回归性能同很多经典学习方法性能相当，比如支持向量机(SVM)。回归参数w有闭合求解公式：

若是在复数域（频域），则其闭合求解公式为：

其中I为单位矩阵，H为Hermitian转置操作算子，即。另外：

上述求解的详细过程如下所示：

基于如下基础求导法则：

对w求导可得：

最终得到：

当样本个数和特征维数都很多时，需对一个**大维数矩阵求逆**，才能获取参数w。在密集样本条件下，那么线性岭回归不满足实时性要求。另外线性岭回归相对于其他非线性学习方法，如非线性核方法，其特征提取表示有限，无法满足稳定鲁棒跟踪对特征表示的要求。**若样本构成循环矩阵，可以通过傅里叶变换将样本矩阵对角化，加快参数求解速度**，从而解决第一个问题。若利用核方法可以在很大程度上解决第二个问题。KCF本质从这两方面展开实现稳定、快速目标平移跟踪。

**循环矩阵：**

给定一个基准向量x，维数为，对其进行循环移位可以构成一个循环矩阵X，维数为。

一维信号的循环矩阵图形表示说明如下。

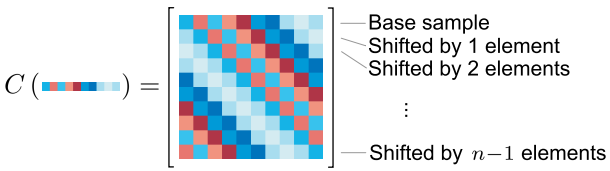


图3 一维信号的循环矩阵图形说明

二维信号循环矩阵块图像举例说明如下。



图4 二维信号循环矩阵块图像举例说明

循环矩阵的数学定义：下面给出循环矩阵的两种数学定义表达方式。第一种X的每一行可以表示为：

循环矩阵的第二种表示：每一元素可以表示为：

为表述方便，所有索引都从0开始。

循环矩阵性质：**对角化性质：**

其中为x的离散傅里叶变换（DFT本质上是一种线性操作）即：

n为x的维数，H Hermitian转置操作，F表示离散傅里叶矩阵，是常量矩阵（可以查询相关资料获取计算方法），与数据矩阵x无关。所有循环矩阵能在同一正交基上投影，投影基为离散傅里叶变换基，系数为离散傅里叶变换系数。

结合循环矩阵可分解为频域对角矩阵性质，对上式解析解进行转化，得到：

同时F是酋矩阵，满足：

所以：

由于对角矩阵是完全对称的，可以按元素操作，得到：

其中的表示原始图像块在频域中的自相关，这个概念来自于信号处理领域。

继续推导，利用对角矩阵的性质，可得：

根据F是酋矩阵，继续简化：

两边同时左乘，得到：

根据F是酋矩阵，继续简化：

根据离散傅里叶变换规则：

在根据对角矩阵性质：

即：

注意：以上除法都是矩阵的按元素相除，计算量较小。

**非线性岭回归：**

核方法通过显性核函数，隐性定义某一特征变换映射函数，将原特征转换为高维特征，以提高判别能力。核方法包含两方面。第一：显性核函数和隐性特征变换函数之间关系如下：

第二：参数w可以表示成输入样本的线性加权组合形式：

其中：

为此目标函数在基本空间关于参数w极小化问题转换为对偶空间关于参数极小化问题。也存在闭合解析解：

核方法内在包含的两个层面，通过隐性的线性或非线性特征映射无需进行显式计算映射特征。不仅提升了判别分类能力，而且将原先基本空间线性参数优化问题也变为对偶空间线性参数优化问题。尽管如此，其存在两个方面问题。**第一同线性岭回归方法一样，需求解一个大矩阵的逆**。矩阵大小同样本个数相同。第二，对偶空间映射函数的评估计算量随样本个数增加而增加式7。而基本空间映射函数则不存在此问题。**若核矩阵为循环矩阵、时，则可以利用傅里叶变换将其对角化，解决上述两个问题。**

**模型求解：**

下面给出KCF运算的三个流程。

1、快速参数回归

下面再次给出对偶空间回归参数闭合解析解公式

若K为循环矩阵，则在DFT域的闭合解析解为。

其中为回归目标y的离散傅里叶变换，为循环核矩阵第一行即循环核矩阵的基本样本的离散傅里叶变换。称为核自相关量。任意两个向量的核相关量的元素定义为：**表征两个向量投影至高维空间的相关程度。**

从上述可以看出若核矩阵为循环矩阵，则能实现快速参数回归。那么当样本集X为循环矩阵时，在什么条件下能保证核矩阵也为循环矩阵？**当核函数为酋不变核函数时，核矩阵K为循环矩阵。**

2、快速检测

跟踪过程中通过参数回归学习后，存储了学习参数和跟踪模板集Z。检测的任务根据采集到的系列样本集X，利用学习的参数和跟踪模板集Z，计算匹配响应最大的位置，从而实现目标位置跟踪。

若样本集和跟踪模板集都为循环矩阵，并且核函数为酋不变函数时，则对应的核矩阵也为循环矩阵。表征了样本集和模板集之间在高维空间相关程度。由核相关量循环平移组成。表征了基本模板样本同基本观测样本的所有平移样本在高维空间相关程度。

所有平移样本集的回归函数值为

取最大值所对应位置为最新目标位置用于后续参数回归阶段选取观测样本。

3、快速核相关量计算

参数回归和检测计算中都需要计算核相关量。参数回归需要计算基本观测样本核自相关量；检测需要计算基本观测样本和基本跟踪模板核相关量。若逐个计算每个核相关量对于一个长度为n的信号量，其复杂度为。对于两种核函数：内积核(如多项式核函数)和径向基核如高斯核函数，利用循环矩阵特性能实现核相关量的快速计算。

**内积核函数：**

内积核函数具有如下形式。核相关量的元素和核相关向量可以表示为：

利用循环矩阵的性质可以获得核相关量快速计算公式：

特别的当核函数为多项式核函数时：

则对应核相关量为：

为此对于一个长度为n的信号量，核相关量为计算复杂度为降为。

**径向基核函数：**

径向基核函数具有如下形式：

核相关量的元素可以表示为：

对于循环平移样本，和只需计算一次，只需获取最后一项的快速计算方式。最后一项同内积核函数具有相同形式。同理可以得出径向基函数核相关量的快速计算方式：

特别的当核函数为高斯核函数时核相关量计算公式如下：

对偶相关滤波器(dual correlation filter, DCF)实际是KCF的一种特殊形式，其核函数为一种特殊的内积核函数既。对偶相关滤波器的求解过程和步骤同KCF一致，这里就不再给出描述。只给出核相关量特定的求解公式为。

**DCF同MOSSE滤波器比较：**

尽管这个公式同MOSSE滤波器相同。但是其同MOSSE有如下几个不同点。MOSSE公式是在频域对目标函数进行极小化求解得到。另外MOSSE中的λ参数没有明确物理意义仅仅防止分母为0，而引起数值求解问题。而此处的系数为岭回归问题的正则项权重。另外原始MOSSE无法直接拓展至多通道图像，若拓展会带来很大计算量。而新MOSSE能直接拓展至多通道图像。关于多通道拓展在多通道拓展章节再做进一步介绍。

虽然新MOSSE和DCF都是在原特征内积空间进行参数求解并且都有快速求解方法，两者判别能力相似，但是新MOSSE求解的是基准空间的参数w而DCF求解的是对偶空间的参数。另外都将两者拓展至多通道图像时，**MOSSE是一个单通道多基本样本参数学习问题，DCF是一个多通道单基本样本参数学习问题**。两者实时性都很高，满足实时跟踪要求。

**多通道处理：**

虽然目前已有几类方法用将线性相关滤波器扩展至多通道，但是其是将多通道分解为多个优化问题求解或分解为其他优化求解问题，计算量大，无法满足实时跟踪要求，只能应用于离线跟踪。**KCF和DCF多通道处理通过将所有通道顺序连接构建成一个基本样本用于构成循环矩阵样本解决计算量大问题**。而新MOSSE公式转换为单通道多基本样本参数学习问题，也能同样满足实时要求。

KCF多通道处理：

KCF利用离散傅里叶变换，在对偶空间进行参数学习。目前KCF采用的核函数都是一向量的点积或范数作为输入参数。因此，此类核函数求取扩展至多通道时只需将各通道的内积累加即可。离散傅里叶变换属于线性变换，其可以在频域对各通道进行累加从而实现多通道处理。对于有C通道的多通道数据，将各通道样本顺序连接形成基本样本用于构成循环矩阵样本。对于多通道KCF，核相关量求解只需按单通道方式将各通道特征累加求和即可。特别地对于高斯核函数核相关量求解公式如下：

DCF多通道处理：

当为DCF时，其对应的多通道核相关函数求解公式如下：

因为在核相关求取过程中已将所有通道量合并完毕，参数学习和检测公式同单通道一致。

MOSSE多通道处理：

MOSSE公式转换为单通道多基本样本参数学习问题。假设有C个通道，每个基本样本可以构成一个循环矩阵。则MOSSE多通道参数w学习公式为：

## 5、实验分析

试验分析：

作者利用OTB提供的50个跟踪视频基准，将基于原始像素的单通道DCF和KCF，基于HOG特征的DCF和KCF同很多主流跟踪方法(TLD，Struct，MIL，MOSSE、ORIA、CT)从性能和运算时间进行了细致比较。这里就不再详述。这里只描述一下个人对这些实验分析的感觉。

MOSSE性能同DCF相当。当用原始像素空间作为输入时，KCF性能明显优于DCF，说明非线性核函数特征提取能力提升了判别能力。若利用HOG特征作为输入时，KCF和DCF性能相当。这说明若已用特征提取算法对目标进行特征提取，KCF在此基础上再进一步可挖掘鲁棒特征能力有限。这里没有比较MOSSE利用HOG进行目标跟踪。我想应该同KCF和DCF性能相当。**通过上面分析可以得出在进行目标跟踪时，特征提取起着十分重要作用。若特征足够好，跟踪方法并不显的那么不重要。**作者在进行算法比较过程中，KCF选取的基本样本窗口大小为目标大小2.5倍。应该**分析一下不同窗口大小对性能的影响**。在性能和运算速度之间选取一个折中。MOSSE由于只选取目标窗口大小进行学习跟踪，所以其每秒可以达到600多帧。然而其性能相对于DCF并没有下降很多。我想窗口大小对KCF和DCF的影响应该不会特别大。

**实际测试分析：**

该实验的测试数据是OTB数据：bolt序列。

实验一：目标框放大系数padding对性能的影响

实验二：更新策略对于跟踪性能的影响

## 6、总结展望

**优缺点：**

KCF实际上是利用核函数对目标进行特征提取，实现目标位置跟踪。虽然其试验说不对缩放尺度敏感，但是其整个过程中目标和基准样本尺寸一直保持不变。KCF能适应小尺度缩放，但是应该无法适于大尺度缩放目标跟踪问题。对于KCF可以考虑增加一个目标尺度估计步骤，以满足大尺度缩放跟踪要求。

KCF没有对目标运动进行预测。若在某一瞬间目标运动比较剧烈，直接以没有预测位置为中心选取样本进行加权相关匹配有可能降低匹配精度甚至匹配失败。

KCF无法应对干扰目标对跟踪目标的干扰。我想这是所有跟踪方法的一个短板。

KCF应该加入一个同TLD类似的机制，以防目标丢失后重新检测跟踪。