上海大学无人艇工程研究院

——环境感知组

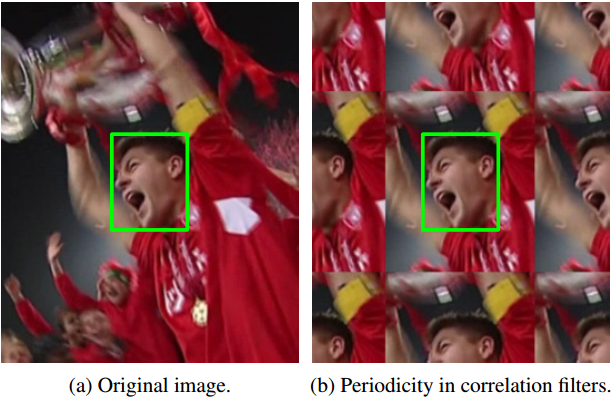
# Learning Spatially Regularized Correlation Filters for Visual Tracking—SRDCF

作者：**Martin Danelljan** Gustav Häger Fahad Shahbaz Khan Michael Felsberg

主页：<http://www.cvl.isy.liu.se/research/objrec/visualtracking/regvistrack/index.html>

出处：2015年ICCV

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：2017.11.29

## 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20日 | 陈加宏 | 完成算法细节的总结 | V1.0.1 |
| 2017年11月29日 | 陈加宏 | 完成算法细节理解的优化 | V1.0.2 |

目 录

[Learning Spatially Regularized Correlation Filters for Visual Tracking—SRDCF 1](#_Toc493764407)

[版本更新记录 2](#_Toc493764408)

[1、概述 3](#_Toc493764409)

[1.1 前言——研究背景及意义、该领域存在的问题 3](#_Toc493764410)

[1.2 创新点——本文算法要解决的问题以及具体的解决方法 5](#_Toc493764411)

[2、细节 5](#_Toc493764412)

[2.1 主要流程 5](#_Toc493764413)

[2.2 数学模型 6](#_Toc493764414)

[2.3 模型求解 7](#_Toc493764415)

[3、实验 8](#_Toc493764416)

[3.1 代码框架 8](#_Toc493764417)

[3.2 实验结果及分析 8](#_Toc493764418)

[3.3 优缺点总结 9](#_Toc493764419)

[3.4 今后工作 9](#_Toc493764420)

[附件一 10](#_Toc493764421)

## 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（相关与卷积操作在数学上类似，只是卷积核旋转了180度）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，算法的计算复杂度大大的降低了，从而实现了该算法极快669FPS的跟踪效果。但是MOSSE的样本采样仍是一种稀疏采样，训练效果一般；同时采样的是线性滤波器（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（在最小二乘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核技巧转换到对偶空间，进一步证明核矩阵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类算法在实时性方面的杰出表现，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，这时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。还和模型**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，就会变得不认识目标。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：

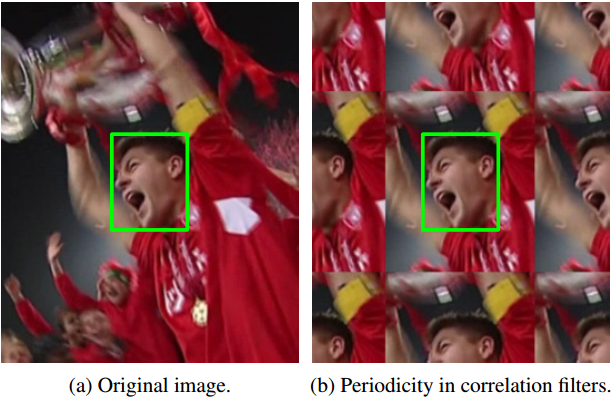


图1 核相关滤波算法中循环假设的示意图

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，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1/4(padding= 1的情况)，即使这样仍然有3/4(30000/40000)的样本是不合理的，这些样本会降低分类器的判别能力。再者，加余弦窗不是免费的，余弦窗将图像块的边缘区域像素全部变成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，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)给传统相关滤波跟踪框架带来的影响，而本文算法试图提升CF类框架中的跟踪鲁棒性，主要从解决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——本文算法要解决的问题以及具体的解决方法

上述的边界问题会影响跟踪器的跟踪性能。带来的影响总结为：

1、由于搜索区域的限制，在目标快速运行时表现不佳；

2、只有少量足够好的负样本容易发生模型的过拟合问题，使模型不能应对目标变形问题；

3、降低了目标被遮挡后再检测的可能性；

4、 直接扩大搜索面积会造成在正样本中包含大量的背景信息，严重降低了模型的判别能力；

解决思路：

1、用空间正则化来减少相关滤波跟踪框架中的周期性假设的不好影响，空间正则项的权重是基于空间范围的先验信息的，这样可以限制滤波器的系数在边缘附近接近0，最大程度的减小边界的影响。由于空间正则化的存在，滤波器可以在更大的范围内进行学习，可以包含更多的负样本，模型的判别能力会增强；

2、为了满足实时性要求，必须提高算法的计算高效性，就提出一种优化策略，主要利用的是空间正则函数在频域内的稀疏性，然后利用高斯-赛德尔模型解决问题，最后可以借助子网格的方法来增大检测得分。

## 2、细节

VOT2015中最亮眼的相关滤波方法当属Martin Danelljan大牛的SRDCF，主要思路：**既然边界效应发生在边界附近，那就忽略所有移位样本的边界部分像素，或者说让边界附近滤波器系数为接近零。**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

这部分说明了MD大神提出该算法的出发点以及具体使用的解决方法。详细的流程关系如下图所示。

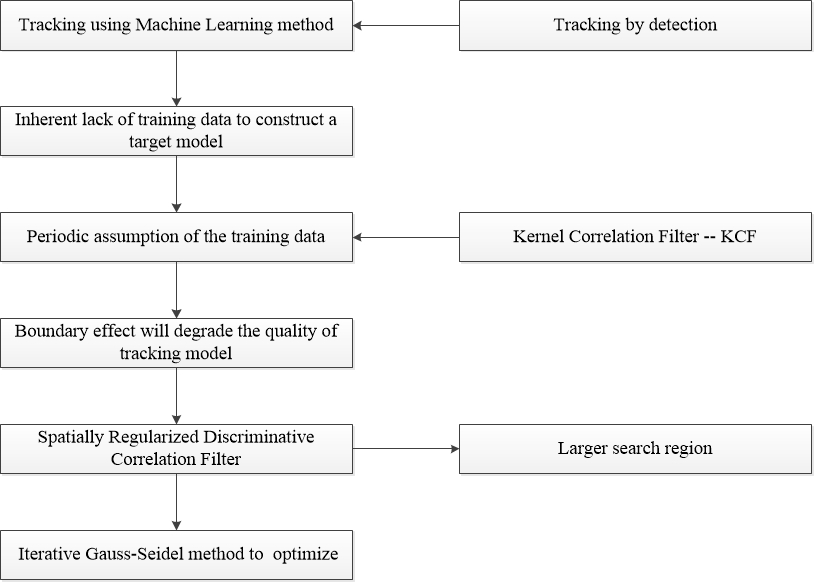


图2 空间正则化算法的基本思路和解决方法示意图

### 2.2 数学模型

本文提出的SRDCF算法是基于DCF框架的，类SAMF多尺度，采用更大的检测区域跟踪参数padding = 4，同时加入空域正则化，惩罚边界区域的滤波器系数。对DCF框架做个简单的介绍， 训练样本多维特征与多个维度的相关滤波器之间卷积后得到的输出响应如下：

根据岭回归模型建立滤波器训练的优化目标函数为：

检测时根据输出响应来找到响应最大的位置就是目标的位置，响应输出为：

这种将时域中的卷积转换成频域中的点乘可以将算法的计算复杂度从原来的变成了，有效提高了算法的执行效率。

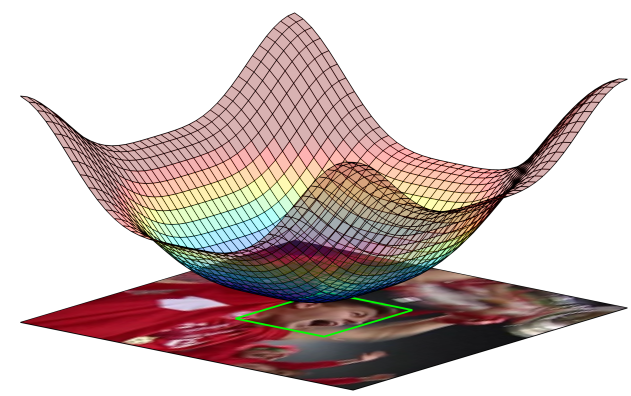


图3 空间正则化算法中对样本中背景信息的惩罚示意图

**SRDCF算法将模型的正则化项的权衡系数从原来的定值改成一种符合空间约束的正则化权重系数，即靠近目标中心的位置受惩罚小，远离目标的位置受到的惩罚大。**那么，原本的正则化项就被更通用的吉尔诺夫正则化项（Tikhonov regularization）代替。这样得到的损失函数变成：

当可以将模型退化成DCF模型，所以上面的加入空间正则化的滤波器训练目标函数比传统的DCF算法更具普适性，但是破坏了原来凸优化的优美结构。

直接将上式使用DFT结合Parseval’s theorem定理：函数平方和等于其傅立叶转换式平方和，得到：

由离散傅立叶反变换中卷积性质可得：

将频域空间正则化项的权衡系数进行归一化处理，可以得到：

对上式进行向量化处理，可以得到：

上述几个优化函数的优化解是一致的，为了实现在实数域的求解，从而保证求解收敛的更快，作者定义了一个大小为MN\*MN的矩阵B使得：，为对应的满足艾尔米特对称性的实数值矩阵，那么由、、可以进一步推导得到：

重新后，可以表示为：

这样整个优化目标就是常见的实数值优化问题，可以直接求导后得到：

这样的求解过程非常耗时，作者根据的稀疏性来用高斯-赛德尔方法对上式进行求解。

上面是我自己根据原文的推导，感觉写的并不够严谨，可以参考网友的较为优秀的推导过程，参见附件一。

### 2.3 模型求解

边界效应发生在原始样本目标中心循环移位到边缘附近时，所以越靠近边缘区域，正则化系数越大，惩罚并抑制这里的滤波器系数(有意忽略边界附近的样本值)。注意区分余弦窗是加在原始图像块上的，之后进行移位产生样本集；而空域正则化是加在样本集中每个移位样本上的。加正则项这种方法属于伤筋动骨类方法，破坏了DCF的封闭解，所以SRDCF采用高斯-塞德尔方法迭代求解最优化，这也是是导致SRDCF比较慢的原因。**padding=4**检测和训练图像块更大，**为了减小计算量SRDCF控制滤波器的大小不超过50\*50**，迭代求解的次数为4，选取的空间正则项分布是一个凸函数，虽然速度只有**5FPS**，但实测对快速运动的提升非常明显。

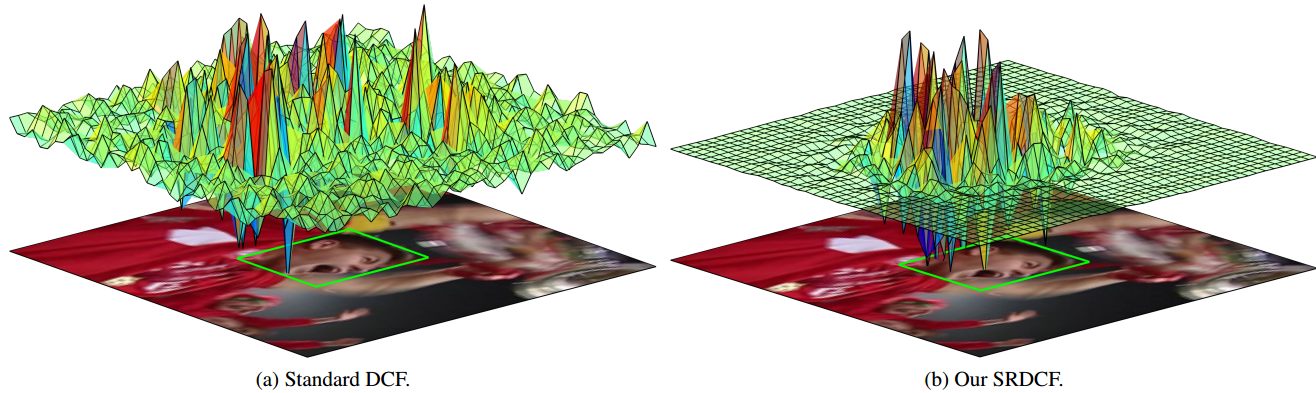


图4 空间正则化后训练得到的空间域中滤波器系数的分布图

整个优化问题由于是满足循环矩阵的，所以采用高斯-赛德尔方法求解，将分解成两个矩阵：下三角矩阵和上三角矩阵，满足：

然后根据迭代求解方法得到：

对上式反复迭代后可求得模型的参数。求得后反向推导后得。

## 3、实验

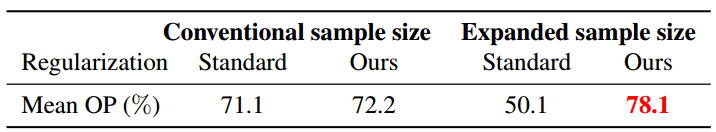
该部分主要讲述算法实现代码的主要流程、实验环境及效果分析、算法优缺点的总结，最后提出后续可改进的方面。实验是检验真理的唯一标准，那么对实验结果详细的分析以及结合算法的原理对算法本质上的一些思考有利于之后研究工作的开展，也是今后工作的一个研究突破点。

### 3.1 代码框架

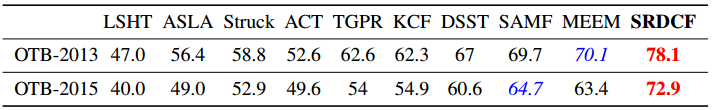


### 3.2 实验结果及分析

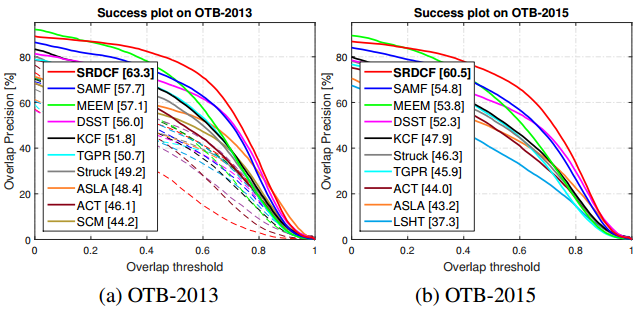
具体的实验效果如下，获得了VOT2015的第四名，其深度特征版本DeepSRDCF是其加强版，是VOT2015的第二名。下表是传统CF算法与SRDCF在padding增大的情况下平均重叠面积成功率的对比，可以看出SRDCF的性能有较大的提升，但是CF类算法的性能大幅度的降低了，也验证了边界效应的影响被减弱后，跟踪性能可以得到提高。

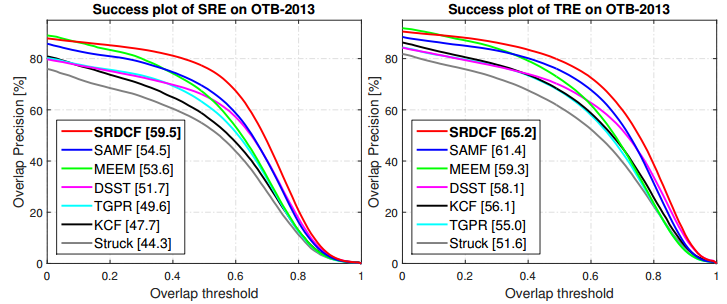


下表是各种优秀算法的成功率的对比，显然在2015年，作为一种只使用了传统特征HOG的SRDCF算法表现是突出的。

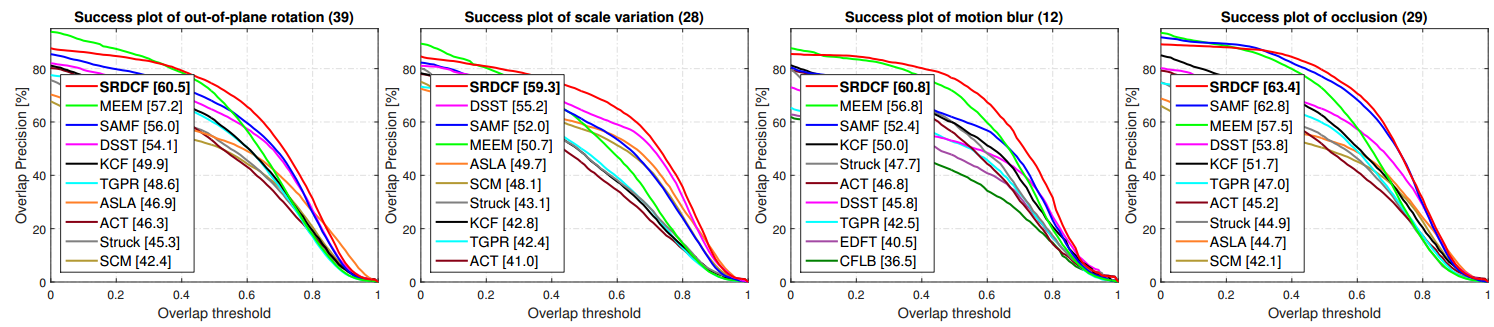


下图是在OTB数据集上详细的的实验结果对比：

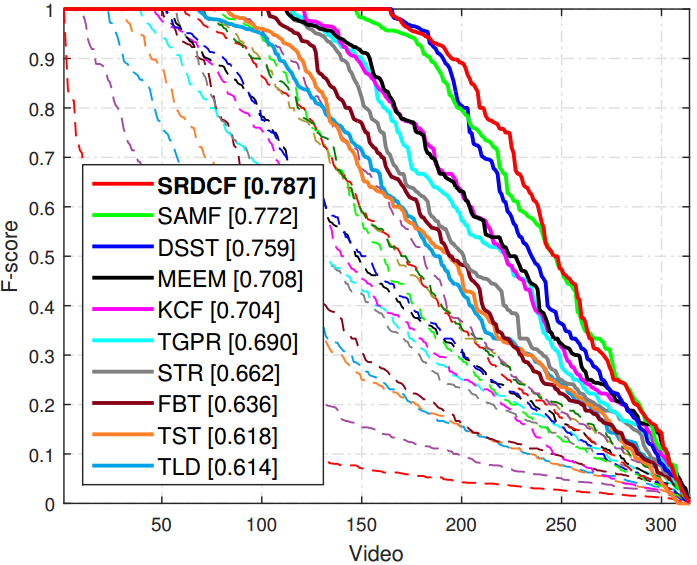


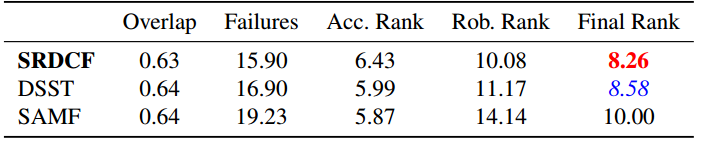


在详细的各类别的数据集上同样表现优异：



作者还在ALOV++和VOT2014数据集上做了实验，实验结果如下所示：





### 3.3 优缺点总结

综合上述的实验以及之前的理论分析，可以总结SRDCF的有点如下：1、在不改变CF整体大框架的前提下减少了边界效应的影响，增强了CF类算法的跟踪鲁棒性，可以更好的应对目标快速运动的情况，也对目标的抖动效果更好；2、提供了一种打破传统CF框架后优化求解的方法。

同时SRDCF算法的缺点也同样明显，速度较慢，在5FPS左右，其中的主要原因是该算法模型的求解是通过迭代优化来实现的，算法的复杂度增加了。

### 3.4 思考总结

针对MD的这篇2015年优秀的跟踪文章，有如下几点值得思考：1、从边界效应出发如何建立的目标优化函数；2、目标函数为何打破了传统CF算法的求解优美性的；3、MD大神又是通过何种优化解法来求解的。只要回答了上述的三个问题，也就理解了本文的核心思想。下面针对上述的三个问题分开论述：1、边界效应对跟踪性能的影响是显著的，所以要提升跟踪性能就必须对其作出相应的措施，本文的思想是在滤波器训练过程中增加了一个对滤波器权重的约束，对靠近目标图像边缘的滤波器系数作出惩罚，让边缘附近的滤波器权重接近于0，这样可以很大程度的减少边界效应的影响，但是看下面的优化函数：

直观感觉只是对第二项的权重加了一个约束，具体是如何打破的CF求解的优美性的来看第二点；2、MD先从一般意义上推导了CF算法的解，可以得到下面的式子：

传统的CF框架的解是这样的，可以看做是上面的一种特例：

传统的框架中满足循环矩阵，可以通过对角化简化求解，但是由于加了空间正则化后的解多了这一项，已经不满足了循环矩阵的要求，自然就不能进行对角化来求解，这就是SRDCF算法破坏了CF类存在优美闭式解的原因；3、那针对上述问题，MD是如何求解的呢？他将分解，然后就可以套入高斯-赛德尔方法中进行迭代求解。

## 附件一

网上关于SRDCF比较合理的求解推导过程，可以作为参考。

