

Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual

创建时间：2017/9/23 18:53

作者：XY.Wang

Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking

传统的DCF模型（如SRDCF）采用handcrafted或CNN等特征，但这些特征都是单一分辨率的，即每个通道的特征图的分辨率都是相同的。为了能够将多种分辨率的特征（例如，卷积网络不同层的特征图，层数越深，特征图越小）整合在一起，C-COT提出将卷积过程转换到一个连续域。这个连续域指的是自变量 t 在 $[0, T)$ 范围内。

下面结合公式来描述：

(1) 连续域的卷积定义：

首先介绍连续域时，即 t 在 $[0, T)$ 范围内时，卷积运算的定义如下：

$$g * h(t) = \frac{1}{T} \int_0^T g(t-s)h(s) ds$$

(2) 假设有 M 个训练样本 $\{x_1, \dots, x_j, \dots, x_M\}$

- 每个样本的空间为 $\mathcal{X} = \mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_D}$ 其中 N_1 是第一通道特征图的分辨率, N_D 是第 D 个通道的空间点个数。
- 每个训练样本有 D 个通道（即有 D 个通道的特征） x_j^1, \dots, x_j^D
- 每个通道的特征表示为（ N_d 是第 d 个通道特征的size） $x_j^d \in \mathbb{R}^{N_d}$

第 d 个通道的
连续域特征图

$$J_d\{x^d\}(t) = \sum_{n=0}^{N_d-1} x^d[n] b_d \left(t - \frac{T}{N_d} n \right).$$

内插公式：

作用：将特征图从离散域转换到连续域，原来多种分辨率的样本全都统一到了相同的定义域 $t \in [0, T)$

示意图：

