

数字图像分析第四次实验报告

院 系： 信息科学技术学院

学 号： SA19023005

姓 名： 景军元

授课教师： 李厚强、周文罡

课程助教： 欧阳剑波、刘一衡、林丰

实验名称： 基于相关滤波的视频目标跟踪

一、基本原理

1、岭回归

岭回归的目标在于学习映射函数 $f(\mathbf{z}) = \mathbf{w}^T \mathbf{z}$ 来最小化如下的均方误差：

$$\min_{\mathbf{w}} \sum_i (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2$$

岭回归具有如下的闭合解：

$$\mathbf{w} = (X^T X + \lambda I)^{-1} X^T \mathbf{y}$$

其中 X 为样本矩阵，每一行代表一个样本 X_i ，上述闭合解的复杂度源于矩阵求逆。

相关滤波器（Correlation Filter），源于岭回归，但通过频域求解以及循环矩阵的性质巧妙避免上述矩阵求逆过程。

2、相关滤波器

在岭回归过程中，我们可以对训练数据进行循环移位式数据扩增。即，对于一个基样本，我们可以生成很多如下的假样本。基样本的标签为 1，其他样本根据移位程度的不同，样本标签逐渐趋于零。

通过将单张图片特征进行循环移位，数据矩阵 X 有如下形式：

$$X = C(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix}$$

我们通过构造上述的数据矩阵，目的在于极大加速岭回归的求解。

将前述的闭合解转化为复数域，如下：

$$\mathbf{w} = (X^H X + \lambda I)^{-1} X^H \mathbf{y}$$

其中 H 代表共轭转置。此外，由于矩阵 X 是循环的，可以通过离散傅里叶变换（DFT）对角化：

$$X = F \text{diag}(\hat{\mathbf{x}}) F^H$$

其中 \wedge 表示傅里叶变换， F 为DFT矩阵， $F^H F = I$ 。

根据如上的公式，可以得到岭回归的频域闭合解：

$$\begin{aligned} \mathbf{w} &= (F \text{diag}(\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}}) F^H + \lambda F^H I F)^{-1} X^H \mathbf{y} \\ &= (F \text{diag}(\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda) F^H)^{-1} X^H \mathbf{y} \\ &= F \text{diag}(\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda)^{-1} F^H F \text{diag}(\hat{\mathbf{x}}) F^H \mathbf{y} \\ &= F \text{diag}\left(\frac{\hat{\mathbf{x}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda}\right) F^H \mathbf{y} \\ \hat{\mathbf{w}} &= \text{diag}\left(\frac{\hat{\mathbf{x}}^*}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda}\right) \hat{\mathbf{y}} = \text{diag}\left(\frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda}\right) \end{aligned}$$

这个公式完全避免了矩阵求逆，元素间的乘积极其高效。

当给出一个新的数据矩阵 Z ，我们可以通过如下预测其回归值：

$$\mathbf{r}_p(\mathbf{w}, Z) = Z \mathbf{w} \Leftrightarrow \hat{\mathbf{r}}_p = \hat{\mathbf{Z}} \odot \hat{\mathbf{w}}$$

二、实验内容

1、问题重述

完成 learn_CF.m 函数。该函数的输入为基样本 \mathbf{x} 及其标签 \mathbf{y} 。按照前面的理论推导，相关滤波器具有闭合解，我们并不需要真正地对基样本进行循环移位，而是样本和标签进行频域操作。

2、函数实现

(1) learn_CF.m

①源代码

```
function model_wf = learn_CF(x, y, lambda)
X = fft2(x);
Y = fft2(y);
model_wf = (conj(X).*Y)./(conj(X).*X+lambda);
end
```

②函数说明

根据公式 $\hat{\mathbf{w}} = \text{diag} \left(\frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda} \right)$ 即可得到该函数。

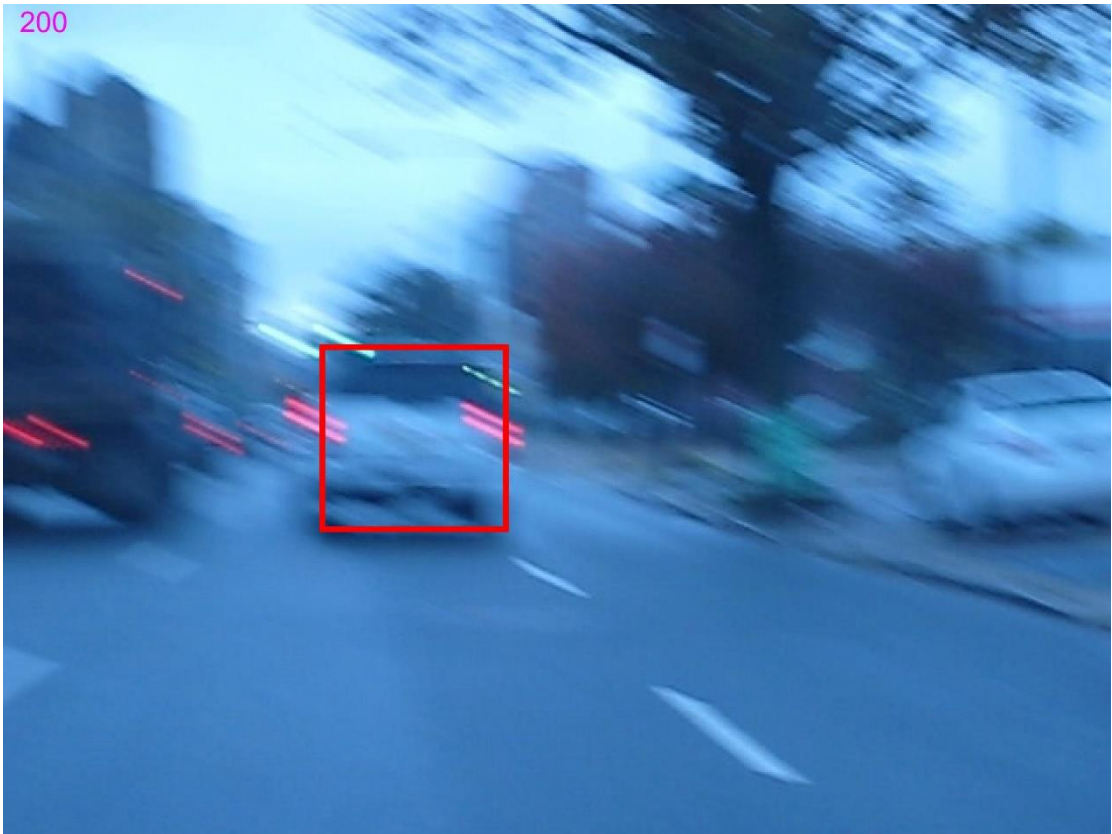
三、实验结果及分析

1、实验结果

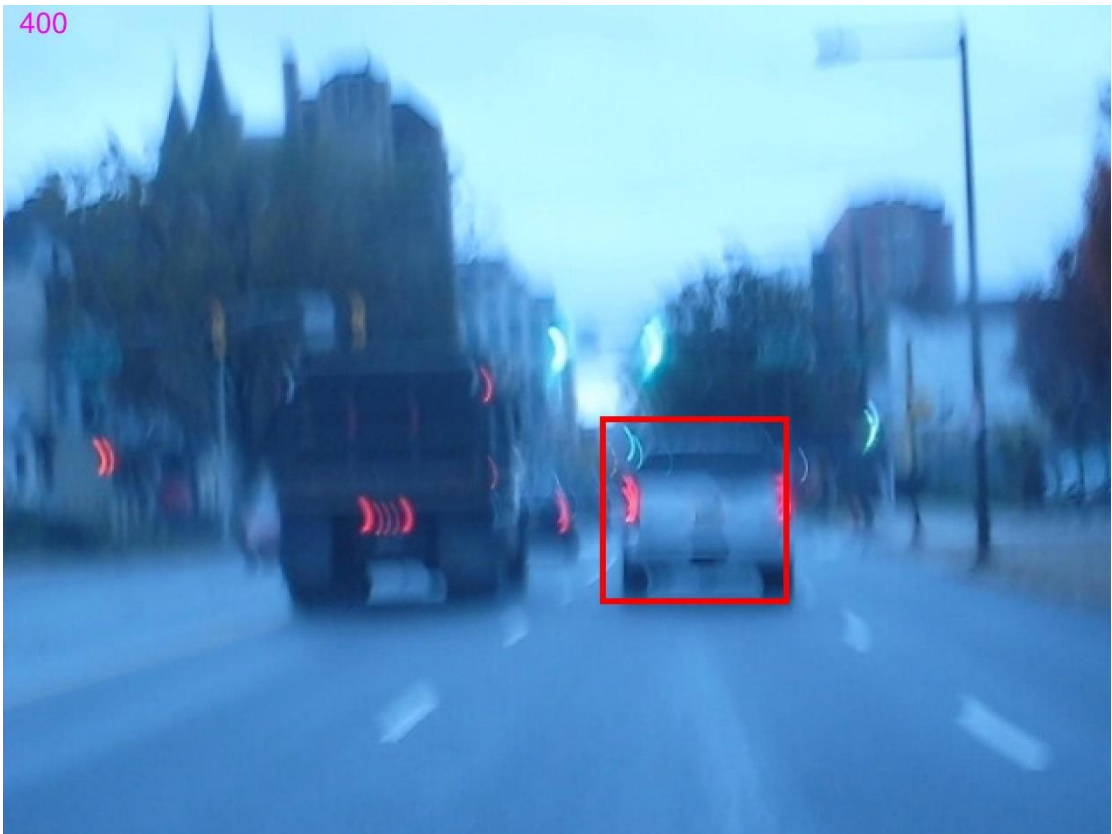
对于 BlurCar1，可以看到第 1 帧匹配结果如下：



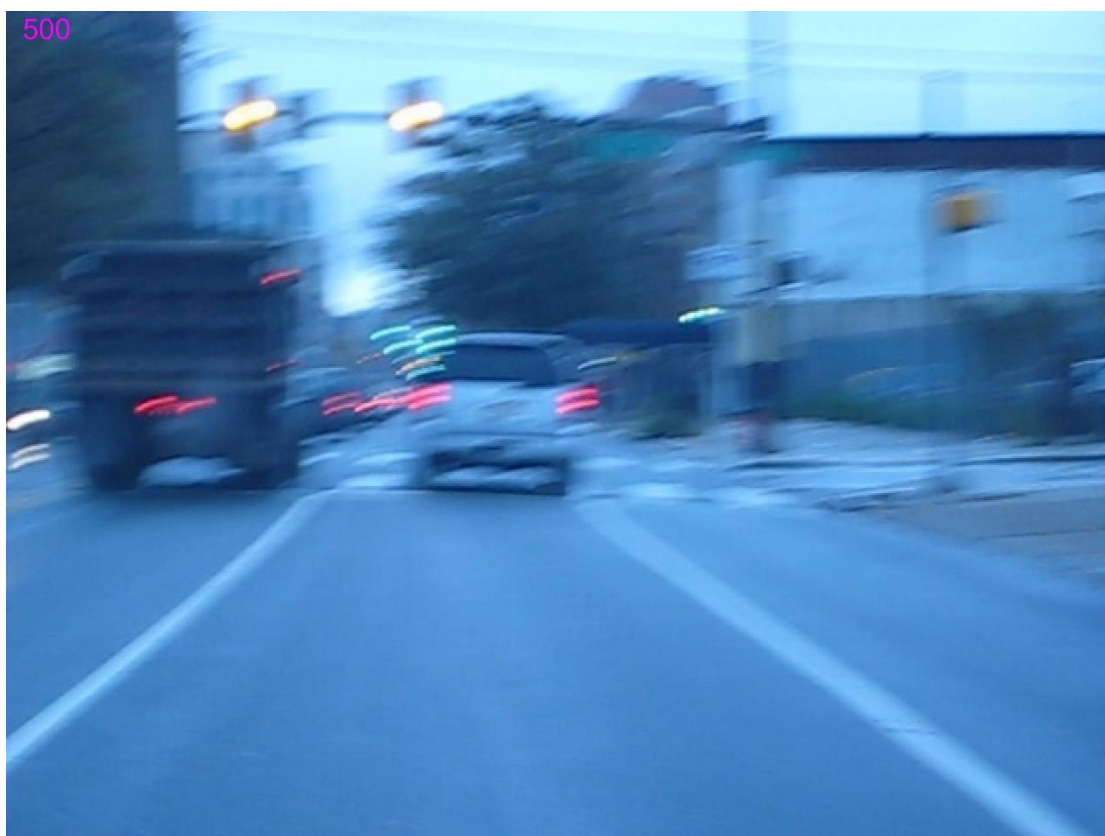
第 200 帧结果：



第 400 帧结果:



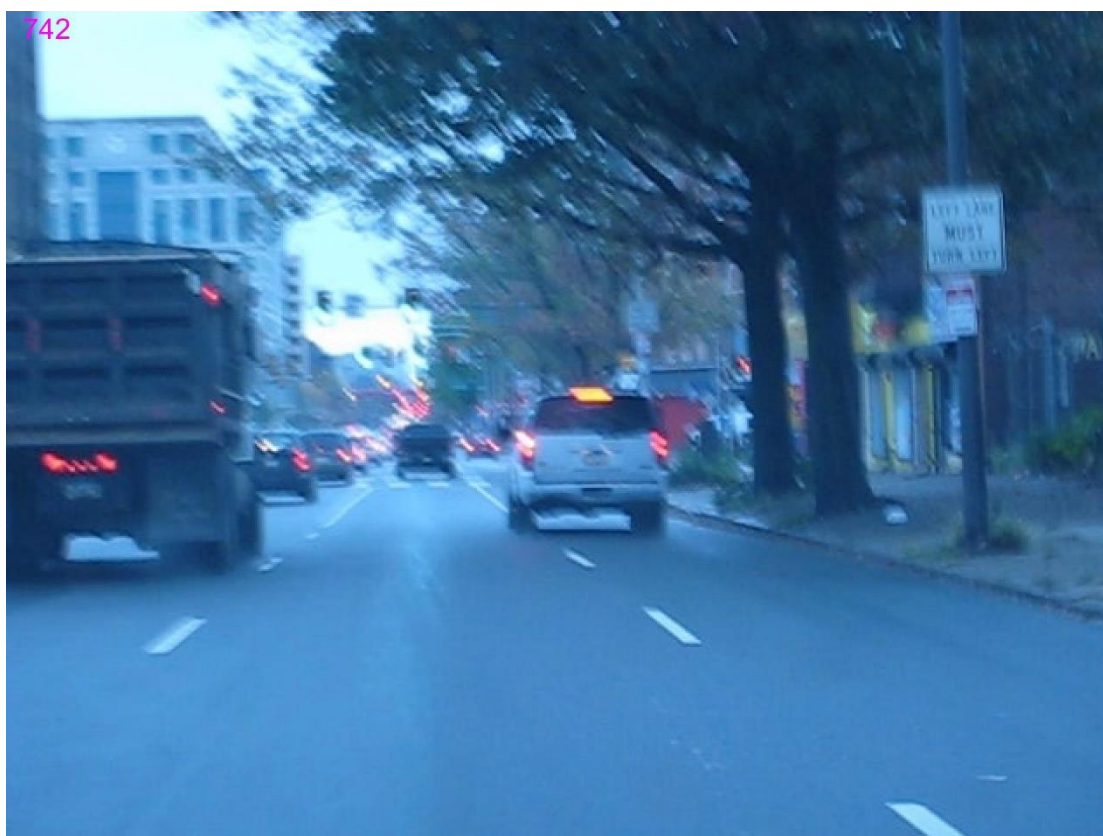
第 500 帧结果:



第 600 帧结果:



第 742 帧（最后一帧）结果:



可以看到，在大约 400-500 帧之间，匹配方法失配。查找失配的位置，可以发现是在 450 帧左右的位置。此时车辆变道，从中间车道变往右侧车道。分析可能是由于汉宁窗边缘带有一定的过渡部分，从而导致周围的一部分区域同样被计为特征图。当车辆变道时，车辆附近的区域有一定的变化，这种变化引起算法失效，进而失配。

通过上述分析，试着将汉宁窗改为矩形窗，结果如下图。虽然没有失配，但是匹配到的位置发生了变化，分析这有可能是由于这种算法对左右方向的变化并不能很好的识别。



如下图是另一个视频的匹配结果，这个视频全程的匹配结果较好，左右的晃动也并没有影响匹配结果，甚至转身之后扔能匹配。分析原因是由于该算法匹配的是灰度特征，而这个视频中视频后方的背景是一个浅色的背景，和前景有较高的对比度，因此能比较好的识别。



四、实验收获

- 1、了解了相关滤波进行目标匹配跟踪的算法步骤。
- 2、分析了算法中存在的一些问题，进一步理解了算法的应用限制。