

2. 计算机视觉中的表示

Prof. : 赵旭 zhaoxu@sjtu.edu.cn

TA: 张烨圣 preacher@sjtu.edu.cn

问题 1: LK 光流的计算

现有连续视频中的图像 I^1 ，其经过 dt 的单位时间后变为图像 I^2 ，这两幅图如下图所示；

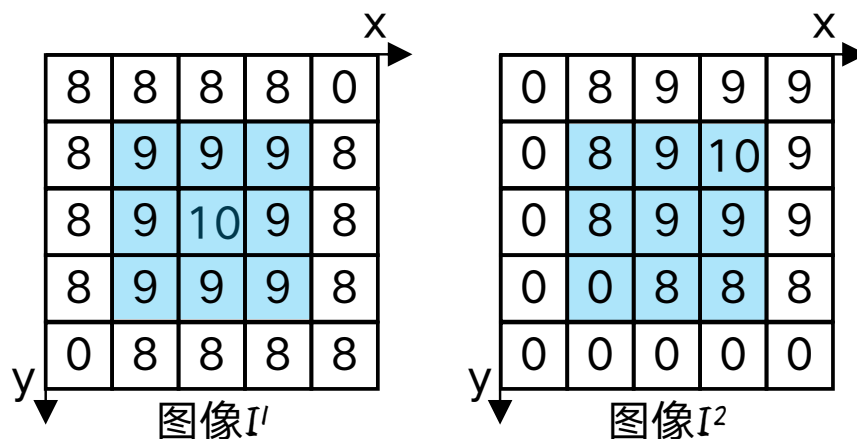


图 1: 问题 1 示意图。

现考虑如图中蓝色阴影部分所示的 3×3 大小的窗口，假设窗口内的像素具有同样的运动；

同时我们约定一种图像梯度和图像关于时间导数的计算方式：

$$\frac{\partial I^1_{x,y}}{\partial x} = \hat{I}^1_{x+1,y} - \hat{I}^1_{x-1,y}; \quad \frac{\partial I^1_{x,y}}{\partial y} = \hat{I}^1_{x,y+1} - \hat{I}^1_{x,y-1}; \quad \frac{\partial I^1_{x,y}}{\partial t} = \hat{I}^2_{x,y} - \hat{I}^1_{x,y},$$

其中， $\hat{I}^1_{x,y}$ 为图像 I^1 在 (x,y) 处的像素值， $\frac{\partial I^1_{x,y}}{\partial x}$ 为图像 I^1 在点 (x,y) 处 x 方向的梯度；

请你用 LK 光流法，计算出像素在 x 轴和 y 轴上的运动速度 $[u,v]^T$ （即 dt 时间内，像素在两个轴上的位移）。

评分标准：共 20 分，根据 LK 光流的计算公式，其中矩阵 A 占 5 分，矩阵 b 占 5 分，位移矩阵 $[u,v]^T$ 占 10 分。

问题 2：实际中的卷积神经网络应用

在常用的深度学习框架 pytorch 中，卷积神经网络的实现类为：

```
1 torch.nn.Conv2d(  
2     in_channels: int, # 输入的通道数;  
3     out_channels: int, # 输出的通道数;  
4     kernel_size: int, # 卷积核大小 (width=height) ;  
5     stride: int, # 卷积步长;  
6     padding: int, # 在输入图像的四周填充 'padding' 层像素;  
7     padding_mode = 'zeros', # 控制填充边缘的像素值, 'zeros'指填充0;  
8     ...  
9 )
```

其中关键的参数注释如上。

假设我们不考虑输入输出的通道数变化，只考虑其长宽；现有输入图片大小为 256×256 ，依次经过一层卷积 (kernel size 5×5 , padding 2, stride 5)，第二层卷积 (kernel size 3×3 , padding 0, stride 1)，第三层卷积 (kernel size 3×3 , padding 1, stride 1) 之后输出。请给出上述每一层卷积后的输出尺寸（只考虑长宽）。

评分标准：共 15 分，每层输出的尺寸占 5 分。

问题 3: SIFT 特征检测与运用

请在 canvas 中下载 ‘Spring24_CV_ug_HW2_3.ipynb’ 文件，并在其中作答。

评分标准：共 65 分，具体见 notebook 文件。