CVPR 第二次作业

强基数学 002　吴天阳　2204210460

1.1　图像参数化几何变换原理

1. 平移变换（自由度为 2）

$$\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{t} \iff \boldsymbol{x}' = \begin{bmatrix} 1 & 0 & t_1 \\ 0 & 1 & t_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

2. 旋转变换（自由度为 1）

$$\boldsymbol{x}' = R_\theta \boldsymbol{x} \iff \boldsymbol{x}' = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

3. 欧式变化（自由度为 3）

$$\boldsymbol{x}' = [R_\theta|\boldsymbol{t}]\boldsymbol{x} \iff \boldsymbol{x}' = \begin{bmatrix} \cos\theta & -\sin\theta & t_1 \\ \sin\theta & \cos\theta & t_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

4. 相似变换（自由度为 4）

$$\boldsymbol{x}' = [R_\theta|\boldsymbol{t}]\boldsymbol{x} \iff \boldsymbol{x}' = \begin{bmatrix} s\cdot\cos\theta & -s\cdot\sin\theta & t_1 \\ s\cdot\sin\theta & s\cdot\cos\theta & t_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

5. 仿射变换（自由度为 6）

$$\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{t} \iff \boldsymbol{x}' = \begin{bmatrix} a_{11} & a_{12} & t_1 \\ a_{21} & a_{22} & t_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

1.2　前向变换与逆向变换

设变换矩阵为 T，原图像记为 $f(\cdot)$，变换后的图像记为 $g(\cdot)$，则有

$$g(T\boldsymbol{x}) = f(\boldsymbol{x}), \qquad g(\boldsymbol{x}) = f(T^{-1}\boldsymbol{x}).$$

其中，前者为前向变换（forward warping），后者为逆向变换（inverse warping）.

前向变换中，由于 $\boldsymbol{x}$ 的参数为整数，而 $T\boldsymbol{x}$ 不一定为整数，所以填充时会出现空缺部分；而逆向变换中，计算 $T^{-1}\boldsymbol{x}$ 非整数时，可通过像素插值算法获得该像素处的近似值，可以很好解决空缺问题.

1.3 下抽样原理与内插方法原理

1.3.1 采样定理

下抽样原理即采样定理. 下面定理描述的是对一维信号进行采样的结论，可类比得到二维图像的采样结论.

定理 1.1 (Shannon-Nyquist 定理，采样定理)**.** 设采样频率为 f_s，信号中最大频率为 f_{max}，当 $f_x > 2f_{max}$ 时，采样后的信息完整保留了原始信号的信息，也即可通过采样信息复原出原始信号.（实际引用中一般取采样频率为最大频率的 $2.56 \sim 4$ 倍）

证明. 考虑对原信号做基数为 N 的离散 Fourier 级数展开，将一组 Fourier 基记为（即对复平面单位圆做 N 等分）

$$B_k(x) = \mathrm{e}^{\frac{2\pi \mathrm{i} kx}{N}} = \mathrm{e}\left(\frac{kx}{N}\right), \quad (0 \leqslant k \leqslant N-1)$$

其中 $\mathrm{e}(x) = \mathrm{e}^{2\pi i x}$. 于是原始信号可表示为

$$g(x) = \sum_{k=0}^{N-1} c_k B_k(x).$$

其中 c_k 可通过对原信号进行快速 Fourier 变换得到.

注：由 Euler 公式可知，$B_k(x) = \cos(\dfrac{2\pi kx}{N}) + \mathrm{i}\sin(\dfrac{2\pi kx}{N})$，所以 B_k 在原信号中对应的频率为 $\dfrac{k}{N}$. 由于单位圆具有对称性，$\forall 0 \leqslant k \leqslant \lfloor \dfrac{N-1}{2} \rfloor$ 有 $B_{N-k} = B_{-k}$，所以 B_k 与 B_{N-k} 具有相同的频率. 这说明，如果 f 为 Fourier 级数中最大的频率，则频域图的周期为 $2f$.

假设采样周期为 P，则采样频率为 $f_s = \dfrac{1}{P}$，则有

$$B_{k+\frac{N}{P}}(nP) = \mathrm{e}\left(\frac{(k+\frac{N}{P})nP}{N}\right) = \mathrm{e}\left(\frac{knP}{N} + n\right) = \mathrm{e}\left(\frac{knP}{N}\right) = B_k(nP), \quad (n = 1, 2, \cdots)$$

上式说明，若原信号中同时存在频率为 $\dfrac{k}{N}$ 与 $\dfrac{k}{N} + \dfrac{1}{P}$ 的信号，则它们会在 $P, 2P, \cdots, nP$ 处取值相同，则无法通过采样信息将这两种频率的区分开，所以只需保证这两种频率的信号不同时出现在原信号中即可.

设 f_{max} 为原信号中的最大频率，且能通过 Fourier 级数表出，由上述的注释可知，频域的周期为 $2f_{max}$，所以

$$\frac{k}{N} + \frac{1}{P} - \frac{k}{N} > 2f_{max} \Rightarrow f_s > 2f_{max}.$$

$\square$

1.3.2 内插方法原理

内插方法：设原图像大小为 $N\times M$，记为 $f(x,y)$, $x\in[1,N], y\in[1,M]$，考虑二维平面中非整数点 (x^*,y^*)，一种求解 $f(x^*,y^*)$ 的方法.（也即将 f 延拓到 $\mathbb{R}^2$ 中）

近邻插值：使用原图中距离 (x^*,y^*) 最近的像素进行代换. 记

$$(x_0,y_0)=\mathop{\arg\min}_{(x,y)\in[1,N]\times[1,M]\cap\mathbb{N}^2}||(x,y)-(x^*,y^*)||_2,$$

其中 $||\cdot||_2$ 表示 2-范数，则 $f(x^*,y^*)=f(x_0,y_0)$.

双线性插值：将 (x^*,y^*) 与周围整数点所围成的面积反比作为整数点对应像素的加权值，设 $O(x^*,y^*)$ 周围存在四个整数点 $A(x_h,y_h)$, $B(x_h,y_l)$, $C(x_l,y_l)$, $D(x_l,y_h)$，对应的面积分别为 $S_A=S(O,C)$, $S_B=S(O,D)$, $S_C=S(O,A)$, $S_D=S(O,B)$，其中 $S(O,A)$ 表示由点 O,A 所围成的面积，参考图1. 则有

$$f(x^*,y^*)=S_Af(A)+S_Bf(B)+S_Cf(C)+S_Df(D).$$

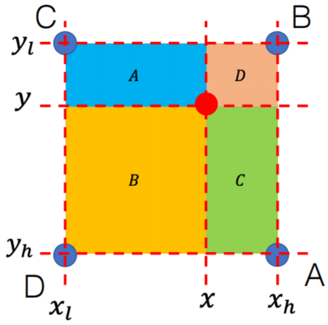

图 1: 双线性插值

1.4 几何变换实验

1.5 Harris 角点检测

设图像为 $I(\cdot;\cdot)$，考虑一个大小为 $2k+1\times 2k+1$ 中心位于 $\boldsymbol{x}_0\in\mathbb{R}^2$ 的滑动窗口 $W_k(\boldsymbol{x}_0)$，任取一个滑动方向 $\boldsymbol{t}\in\mathbb{R}^2$，定义在 $\boldsymbol{t}$ 方向上的**平方误差和（sum of squared differences, SSD）**为

$$E_{\boldsymbol{x}_0}(\boldsymbol{t})=\sum_{\boldsymbol{x}\in W_k(\boldsymbol{x}_0)}(I(\boldsymbol{x}+\boldsymbol{t})-I(\boldsymbol{x}))^2$$

若 $\boldsymbol{x}_0$ 为角点，则 $\forall\boldsymbol{t}\in\mathbb{R}^2$，$E_{\boldsymbol{x}_0}(\boldsymbol{t})$ 都应尽可能大.

记 $\boldsymbol{t}=(u,v)^T$，由 Taylor 公式可知

$$I(\boldsymbol{x}+\boldsymbol{t})=I(\boldsymbol{x})+I_x(\boldsymbol{x})u+I_y(\boldsymbol{x})v+O(||\boldsymbol{t}||_2^2)$$

于是可对 $E_{\boldsymbol{x}_0}(\boldsymbol{t})$ 进行进一步分解

$$\begin{aligned}E_{\boldsymbol{x}_0}(\boldsymbol{t})&\approx\sum_{\boldsymbol{x}\in W_k(\boldsymbol{x}_0)}(I(\boldsymbol{x})+I_x(\boldsymbol{x})u+I_y(\boldsymbol{x})v-I(\boldsymbol{x}))^2\\&=\sum_{\boldsymbol{x}\in W_k(\boldsymbol{x}_0)}(I_x(\boldsymbol{x})u+I_y(\boldsymbol{x})v)^2.\end{aligned}$$

更一般的，设图像 I 大小为 $M\times N$，记全体像素点集为 $D:=[1,M]\times[1,N]\cap\mathbb{Z}^2$，则

$$
\begin{aligned}
E_{\boldsymbol{x}_0}(\boldsymbol{t}) &\approx \sum_{\boldsymbol{x}\in D} w_{\boldsymbol{x}_0}(\boldsymbol{x})\left(I_x(\boldsymbol{x})u+I_y(\boldsymbol{x})v\right)^2\\
&=\sum_{\boldsymbol{x}\in D} w_{\boldsymbol{x}_0}(\boldsymbol{x})I_x^2(\boldsymbol{x})u^2+2w_{\boldsymbol{x}_0}(\boldsymbol{x})I_x(\boldsymbol{x})I_y(\boldsymbol{x})uv+w_{\boldsymbol{x}_0}(\boldsymbol{x})I_y^2(\boldsymbol{x})v^2\\
&=A(\boldsymbol{x}_0)u^2+2B(\boldsymbol{x}_0)uv+C(\boldsymbol{x}_0)v^2.
\end{aligned}
$$

其中 $w_{\boldsymbol{x_0}}(\boldsymbol{x})=\begin{cases}1, & \boldsymbol{x}\in W_k(\boldsymbol{x}_0),\\ 0, & \texttt{otherwise}.\end{cases}$ 为 $W_k(\boldsymbol{x}_0)$ 的示性函数，为使其变化更平滑，可取 $w_{\boldsymbol{x}_0}$ 为中心在 $\boldsymbol{x}_0$ 大小为 $2k+1\times 2k+1$ 的 Gauss 核，上式中 A,B,C 定义如下，分别表示将核 w 作用在 I_x^2, I_xI_y, I_y^2 图像上得到的结果

$$
A(\boldsymbol{x_0})=\sum_{x\in D} w_{\boldsymbol{x_0}}(\boldsymbol{x})I_x^2(\boldsymbol{x}),\ B(\boldsymbol{x_0})=\sum_{x\in D} w_{\boldsymbol{x}_0}(\boldsymbol{x})I_x(\boldsymbol{x})I_y(\boldsymbol{x}),\ C(\boldsymbol{x_0})=\sum_{x\in D} w_{\boldsymbol{x}_0}(\boldsymbol{x})I_y^2(\boldsymbol{x}).
$$

进一步，可使用二次型矩阵表出

$$
E_{\boldsymbol{x}_0}(\boldsymbol{t})=\boldsymbol{t}^TM_{\boldsymbol{x}_0}\boldsymbol{t},\quad \text{其中}M_{\boldsymbol{x_0}}=\begin{bmatrix}A(\boldsymbol{x_0}) & B(\boldsymbol{x}_0)\\ B(\boldsymbol{x_0}) & C(\boldsymbol{x}_0)\end{bmatrix}
$$

称 $M_{\boldsymbol{x}_0}$ 为图像在 $\boldsymbol{x}_0$ 处的二阶矩矩阵.

假设 $M_{\boldsymbol{x}_0}$ 的秩为 2，则 $M_{\boldsymbol{x}_0}$ 存在 2 个特征值，记其中较大者为 λ_{max}，较小者为 λ_{min}，对应的特征向量分别为 $\boldsymbol{x}_{max}$, $\boldsymbol{x}_{min}$，由特征向量定义可知，$M_{\boldsymbol{x}_0}\boldsymbol{x}_{max}=\lambda_{max}\boldsymbol{x}_{max}$ 说明窗口 $W_k(\boldsymbol{x}_0)$ 在沿着 $\boldsymbol{x}_{max}$ 方向上移动单位长度，$E_{\boldsymbol{x_0}}(\boldsymbol{t})$ 可达到最大值. 下面对 λ 的大小进行分类讨论

1. 当 λ_{max} 较小时，$E_{\boldsymbol{x_0}}(\boldsymbol{t})$ 沿各个方向变化都较小，则 $\boldsymbol{x}_0$ 处于图像内部平滑区域.
2. 当 λ_{max} 较大且 $\lambda_{max}\approx\lambda_{min}$ 时，$E_{\boldsymbol{x_0}}(\boldsymbol{t})$ 沿各个方向变化都较大，则 $\boldsymbol{x}_0$ 是角点.
3. 当 λ_{max} 较大且 $\lambda_{max}>>\lambda_{min}$ 时，$E_{\boldsymbol{x_0}}(\boldsymbol{t})$ 仅沿 $\boldsymbol{x}_{max}$ 方向变化较大，则 $\boldsymbol{x}_0$ 是边界点.

通过引入**角点响应函数（Corner response function）**用于判断角点

$$
\begin{aligned}
R(\boldsymbol{x}_0)=&\ \det(M_{\boldsymbol{x}_0})-\alpha\cdot\operatorname{trace}(M_{\boldsymbol{x}_0})^2=\lambda_1\lambda_2-\alpha(\lambda_1+\lambda_2)^2\\
=&\ A(\boldsymbol{x_0})C(\boldsymbol{x_0})-B^2(\boldsymbol{x_0})-\alpha(A(\boldsymbol{x_0})+C(\boldsymbol{x_0}))
\end{aligned}
$$

其中 $\alpha\in[0.04,0.06]$ 为超参数. 当 $R(\boldsymbol{x}_0)$ 大于设定阈值时，则判定 $\boldsymbol{x}_0$ 为角点.

最后，再使用**非局部极大值抑制（Non-maxima suppression, NMS）**，若 $\boldsymbol{x}_0$ 是其邻域内的最大值，则对其进行保留，否则删去该点. 通过设定邻域大小，可对角点密度进行调整.

整个 Harris 角点检测算法中，总共存在 4 个超参数，分别为窗口大小 k，响应函数中 α，响应阈值和 NMS 邻域大小.