图像的全景拼接

图像全景拼接方案流程

1. 特征点检测
2. 创建特征点描述子
3. 特征点匹配
4. 几何变换估计
5. 坐标变换

线性几何变换

线性几何变换的定义：$T(x) = Hx$，其中 x 是 n 维坐标，H 是一个可逆矩阵

旋转变换

旋转变换表示平面上一点绕原点逆时针旋转：$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

在 $a = Ra'$ 中，"R 表达一个保持方向的旋转变换"等价于"R 为正交阵且行列式为 1"。

若 R 是一个行列式为 -1 的正交矩阵，则 R 表达的是一个平面旋转变换复合一个反射变换。

平面旋转变换的自由度为 1。

齐次坐标

齐次坐标用一个三维向量 $(x_1, x_2, x_3)^T$ 来表达二维平面上的点。

如果 $x_3 = 0$，说明这个点为一个无穷远点；如果 $x_3 \neq 0$，说明该点为一个正常点。

一个点的齐次坐标的形式不具有唯一性，$(x_1, x_2, x_3)^T$ 和 $k(x_1, x_2, x_3)^T$ 表示的是同一个点。但规范化齐次坐标 $(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1)^T$ 是唯一的。

正常坐标与齐次坐标的转换：$(x_1, x_2)^T \Leftrightarrow k(x_1, x_2, 1)^T$。

规范齐次坐标的旋转变换：$H = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$。

欧氏变换

欧氏变换由旋转和平移复合而成（不考虑反射变换）：$H = \begin{bmatrix} \cos\theta & -\sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$。

平面欧氏变换有 3 个自由度。

相似变换

相似变换由欧氏变换复合一个缩放得来：$H = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$。

平面相似变换有 4 个自由度。

仿射变换

在相似变换的基础上对矩阵左上 2×2 的子矩阵 R 放宽限制为非奇异矩阵，得到仿射变换：$H = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$。

矩阵 A 的 4 个元素相互独立，带来了 4 个自由度：对 A 奇异值分解 $A = UDV^T$，U 和 D 为正交阵，相当于旋转变换，且角度相反；而 D 相当于缩放，两个方向上的缩放系数分别为 A 的两个奇异值。

矩阵 H 左上角的 2×2 子矩阵 R 的行列式 $\det(R)$ 决定了面积的变化以及是否保持方向。当 $\det(R) > 0$ 时，仿射变换保持方向；当 $\det(R) < 0$ 时，仿射变换改变方向。

平面仿射变换有 6 个自由度。

射影变换

仿射变换矩阵的第三行一定为 $(0,0,1)$。而若仅要求 H 为非奇异矩阵，就得到射影变换：$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$。

射影变换可以定义在无穷远点上，而其他变换都只考虑正常点。

由于齐次坐标的不唯一性，射影变换有 8 个自由度。

特征点检测与匹配

哈里斯角点检测

对于图像上某点，以它为中心取一个窗口 W，窗口移动小量 $(\Delta x, \Delta y)$ 之后，新旧窗口的像素值差异为

$$s_W(\Delta x, \Delta y) = \sum_{(x_i, y_i)\in W} (f(x_i, y_i) - f(x_i + \Delta x, y_i + \Delta y))^2$$

进行一阶泰勒近似后得到

$$s_W(\Delta x, \Delta y) = (\Delta x, \Delta y) M (\Delta x, \Delta y)^T$$

其中

$$M=\begin{pmatrix}\sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial x}|_{(x_i,y_i)}\right)^2 & \sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial x}|_{(x_i,y_i)}\right)\left(\frac{\partial f}{\partial y}|_{(x_i,y_i)}\right)\\ \sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial x}|_{(x_i,y_i)}\right)\left(\frac{\partial f}{\partial y}|_{(x_i,y_i)}\right) & \sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial y}|_{(x_i,y_i)}\right)^2\end{pmatrix}$$

可以证明 M 为半正定矩阵，且大部分情况下为正定矩阵。

当 M 为正定矩阵时，方程 $s_W(\Delta x,\Delta y)=1$ 中 $(\Delta x,\Delta y)$ 的轨迹为一个椭圆，椭圆的半长轴长和半短轴长分别为 $\lambda_1^{-0.5}$ 和 $\lambda_2^{-0.5}$。因此越小越圆的椭圆越可能对应一个角点。

在实现时，特征值分解的计算代价较高，通常使用角点程度经验公式：$r(x)=det(M(x))-k(trace(M(x)))^2$。其中 k 为超参数，通常为 0.04 到 0.06。

选择角点不仅需要看 $r(x)$ 是否大于某个阈值，还需要进行非极大值抑制，以保证选出的角点是稀疏的。

哈里斯角点检测算法具有旋转不变性和整体光照不变性，但不具有尺度不变性，原因在于窗口大小是预先设定的，无法自动自适应。

特征描述子

对于图像上一点 x，以其为中心取一个 $s\times s$（提前设定好的值）的窗口，拉成一个列向量并进行单位化，得到的向量 d 称为块描述子，是最简单的特征描述子。

两个描述子之间的常用距离计算方式有 SSD、SAD、NCC

- 平方差之和：$SSD_{dist}(d_1,d_2)=\|d_1-d_2\|_2^2$
- 绝对差之和：$SAD_{dist}(d_1,d_2)=\|d_1-d_2\|_1^2$
- 规范化互相关距离与皮尔逊互相关系数有关：$NCC_{dist}(d_1,d_2)=1-\frac{1}{n}\frac{(d_1-\mu(d_1))\cdot(d_2-\mu(d_2))}{std(d_1)std(d_2)}\in[0,2]$
- 另一种规范化互相关距离定义：$NCC_{dist}(d_1,d_2)=\arccos\left(\frac{1}{n}\frac{(d_1-\mu(d_1))\cdot(d_2-\mu(d_2))}{std(d_1)std(d_2)}\right)\in[0,\pi]$

SIFT 特征点检测

流程：构建高斯尺度空间 > 构建 DoG 尺度空间 > 检测 DoG 极值点 > 粗略极值点位置精化与 DoG 值估计（过滤非特征点） > 计算海森矩阵（过滤边缘点） > 得到特征点的空间位置和特征尺度

SIFT 中的特征点（斑点）不仅具有位置特征，还具有大小特征。

检测斑点使用尺度归一化 LoG 算子：

$$\sigma^2\nabla^2 g=\frac{x^2+y^2-2\sigma^2}{2\pi\sigma^4}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

对尺度大小参数 σ 进行取值可以得到一系列不同尺度大小的尺度归一化 LoG 算子，用这些算子与图像进行卷积并堆叠可以得到尺度空间。图像上每一点 x 在尺度空间中都有一组响应值 $\{r_i\}$，使得 r 取得极值的唯一 σ 称为斑点的特征尺度。

在实现时，为了简洁高效，可以用 DoG 算子浸提替代尺度归一化 LoG 算子：

$$DoG(\sigma) = g(x, y; k\sigma) - g(x, y; \sigma)$$

这样由于卷积运算性质，卷积输出 $DoG(\sigma) * I$ 就是 $g(x, y; k\sigma) * I - g(x, y; \sigma) * I$。

当 $k \to 1$ 时，$DoG(\sigma) \approx (k-1)\sigma^2 \nabla^2 g(x, y; \sigma)$。但即时在 k 较大时，使用 DoG 算子的性能也不会受到明显影响。

在尺度空间中，如果一点的响应值比 26 个邻近值都大或都小，则得到一个候选特征点。特征点取得的极值的是极大值还是极小值与斑点的亮暗结构有关。

该算法的斑点特征尺度会随着图像的缩放而等比变化，具有尺度协变性。同时，该算法具有旋转不变性、光照不变性。

SIFT 特征描述子

SIFT 特征描述子的构建在高斯尺度空间层 g_l 上进行，在 g_l 上与特征尺度 σ 对应的高斯标准差记为 σ_1。

为了使描述子具有旋转不变性，需要确定图像邻域的主方向，并根据主方向进行方向归一化：在 DoG 空间尺度层上划定一个 $9\sigma_1 \times 9\sigma_1$ 的区域 P。对于 P 上的每一点计算出其梯度，构建梯度直方图，并在直方图中选择主方向，然后进行旋转。

特征描述子的构建：将邻域划分为 4×4 个子区域，在每个子区域中构建一个 8 维的梯度方向直方图，最后将所有直方图连接起来形成一个 128 维的向量。实际实现时，每个点的贡献需要进行基于到中心距离的高斯加权。

一张图像产生的所有信息：特征点坐标、特征尺度、斑点主方向、特征描述子

射影矩阵估计

矩阵微分结论

$$\frac{d\mathbf{a}^T\mathbf{x}}{d\mathbf{x}} = \frac{d\mathbf{x}^T\mathbf{a}}{d\mathbf{x}} = \mathbf{a}$$

$$\frac{d\mathbf{x}^T\mathbf{x}}{d\mathbf{x}} = 2\mathbf{x}$$

$$\frac{d\mathbf{y}^T(\mathbf{x})}{d\mathbf{x}} = \left(\frac{d\mathbf{y}(\mathbf{x})}{d\mathbf{x}^T}\right)^T$$

$$\frac{d\mathbf{A}\mathbf{x}}{d\mathbf{x}^T} = \mathbf{A}$$

$$\frac{d\mathbf{x}^T\mathbf{A}^T}{d\mathbf{x}} = \mathbf{A}^T$$

$$\frac{d\mathbf{x}^T\mathbf{A}\mathbf{x}}{d\mathbf{x}} = (\mathbf{A} + \mathbf{A}^T)\mathbf{x}$$

$$\frac{d\mathbf{a}^T\mathbf{X}\mathbf{b}}{d\mathbf{X}} = \mathbf{a}\mathbf{b}^T$$

$$\frac{d\mathbf{a}^T\mathbf{X}^T\mathbf{b}}{d\mathbf{X}} = \mathbf{b}\mathbf{a}^T$$

$$\frac{d(tr\mathbf{X}\mathbf{B})}{d\mathbf{X}} = \mathbf{B}^T$$

$$\frac{d|\mathbf{X}|}{d\mathbf{X}} = |\mathbf{X}|(\mathbf{X}^{-1})^T$$

拉格朗日乘子

用于寻找 $f(x)$ 在等式约束 $g_i(x) = 0$ 下的所有极值点。

定义 $F(x; \lambda_1, \cdots \lambda_m) = f(x) + \sum_i \lambda_i g_i(x)$，则 $f(x)$ 的所有极值点一定都是 F 的驻点。

只需要求解 $n + m$ 个方程并从解中寻找答案。

线性最小二乘

非齐次线性方程组：求 $\arg\min_x \|Ax - b\|_2^2$。由于 $\|Ax - b\|_2^2$ 是 x 的凸函数，驻点 $x_s = (A^TA)^{-1}A^Tb$ 即为全局最小值点。

齐次线性方程组：求 $\arg\min_x \|Ax\|_2^2$，其中 $\|x\|_2 = 1$。令 $L(x, \lambda) = \|Ax\|_2^2 + \lambda(1 - \|x\|_2^2)$，驻点满足 $A^TAx_0 = \lambda_0 x_0, x_0^T x_0 = 1$，即驻点为 A^TA 的所有特征值。其中显然最小值点为 $\min\{\lambda_i\}$。

RANSAC 单应估计

一个射影变换的单应矩阵只需要 4 个方程就能确定，而实际的特征点对数远大于 4。

由于存在异常点（离群值），直接使用最小二乘会受到较大影响。

RANSAC：每一次迭代过程中随机选择所有样本点的一个子集，并进行模型拟合，统计该模型下误差值小于某个阈值的样本点，称为当前模型的”一致集”。重复这一过程，直到次数到达上限或某次一致集的大小大于某个阈值。