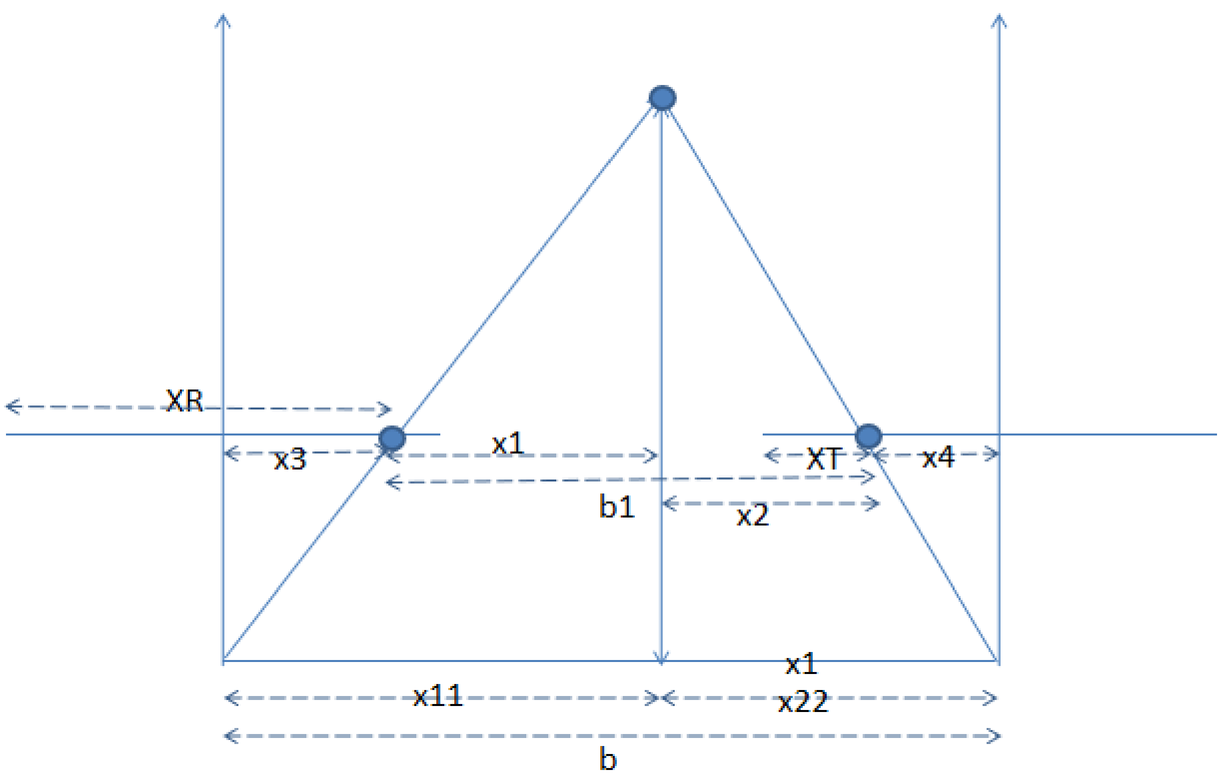
三维视觉与计算摄像学

第2次作业

王信超 2201212814

第1题：

深度与视差之间的关系



深度：深度是空间中某点成像在相机中相机坐标系的Z坐标值，是一个**空间单位**,也就是图中的长度Z.

视差： 视差 d 等于同名点对在左视图的列坐标减去在右视图上的列坐标，是**像素单位**,也就是$d = XR - XT$的长度

深度Z和**视差d**成反比关系

推导过程

由于 $x_1 + x_2 = b_1, x_{11} + x_{22} = b$

由于三角形相似可以得知$\frac{x_{11}}{Z} = \frac{x_1}{Z-f}, \frac{x_{22}}{Z} = \frac{x_2}{Z-f}$

两个式左右相加：可以得到 $\frac{x_{11}+x_{22}}{Z} = \frac{x_1+x_2}{Z-f} = \frac{b_1}{Z-f}$即为 $\frac{b}{Z} = \frac{b_1}{Z-f}$ **(*)**

其中$b_1 = b - x_3 - x_4 = b - (XR - \frac{l}{2}) - (\frac{l}{2} - XT) = b - XR + XT = (b + XT) - XR$

将b_1带入**(*)**式可以得到 $\frac{b}{Z}=\frac{b_1}{Z-f}=\frac{(b+XT)-XR}{Z-f}$

所以$b(Z-f)=Z(b+xT-XR) ==> bf=z(XR-XT)=zd$

所以$Z=\frac{bf}{d}$

所以容易得知：深度Z和视差d成反比关系

关系

深度估计：，就是利用一张或者唯一/多个视角下的RGB图像，估计图像中每个像素相对拍摄源的距离。由于视差和深度之间的反比关系，可以通过根据在两张图像中的视差从而估计像素与相机之间的距离，也就是深度。

视点合成:视点合成是根据深度和视差之间的关系，从而不断改变视差信息，可以获得不同的深度图，最终将深度图每一个像素点映射到三维空间，然后再将每一像素点重映射到对应的虚拟视点上 。

SFM(Structure from Motion):用来估计相机参数和三维点的位置，我们根据深度和视差之间的关系，从而采集匹配的特征点，然后求解相机的参数信息，最后可以用来计算三维点的位置信息。

第2题：

预知识

设三维空间的特征点在**相机坐标系**下的坐标为 $\vec{X}=\begin{pmatrix}x\\y\\z\end{pmatrix}$,

而相机坐标系投影到**像素坐标系的齐次坐标** $\vec{p}=\begin{pmatrix}u\\v\\1\end{pmatrix}=\frac{1}{z}\mathbf{K}\vec{X}$，那么反之由像素齐次坐标求相机系坐标为

$\vec{X}=z\mathbf{K}^{-1}\vec{p}$ **(1)**

- q为像素系齐次坐标，
- K为相机的内参矩阵，其中$\mathbf{K}=\begin{pmatrix}f_x & 0 & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{pmatrix}$

3D空间点p所在的平面在相机坐标系下的平面参数为$\{\vec{n},d\}$，其中$\vec{n}$为平面法向量，d为相机系原点到平面的距离。

那么3D空间点p位于平面$\{\vec{n},d\}$上由方程表达式 $\vec{n}^T\cdot\vec{X}+d=0$,那么带入(1)式可得 $\vec{n}^T\cdot z\mathbf{K}^{-1}\vec{p}+d=0$

那么 $z=-\frac{d}{\vec{n}^T\mathbf{K}^{-1}\vec{p}}$ **(*)**

那么如果我们知道3-D平面的相关参数$\{\vec{n},d\}$，我们就能够求出3-D点的深度z,从而结合(1)式由像素齐次坐标 $\vec{p}$ 得出相机坐标 $\vec{X}=-\frac{d}{\vec{n}^T\cdot\mathbf{K}^{-1}\vec{p}}\cdot\mathbf{K}^{-1}\vec{p}$

单应性变换矩阵H

通过单应性矩阵的映射变换，可以将**一副图像上的点**转换成为**另一副图像上的点**，单应性变换矩阵H 是两幅图像上的**像素坐标**转换的变换矩阵。

单应性矩阵的推导过程

相机坐标系a下的点 $\mathbf{X_a}$ 和相机坐标系b下的点坐标$\mathbf{X_b}$ 其中$\mathbf{X_a},\mathbf{X_b}$是相机坐标系下的三维坐标，等同于 $\mathbf{X_a} = \vec{X}_a = \begin{pmatrix} x_a \\ y_a \\ z_a \end{pmatrix}$, $\mathbf{X_b}$ 类似于$\mathbf{X_a}$.

根据相机成像原理，相机坐标系a下的点 $\mathbf{X_a}$转换成为相机坐标系b下的点坐标$\mathbf{X_b}$ 公式如下：

$$\begin{pmatrix} \mathbf{X_b} \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R_{ba}} & \mathbf{t_{ba}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{X_a} \\ 1 \end{pmatrix}$$

$\mathbf{R_{ba}}$→表示 a系 到 b系的坐标旋转变换(也可以理解为，b系下a系的姿态)

$\mathbf{t_{ba}}$ →表示 a系 到 b系的坐标平移变换(也可以理解为，b系下a系的位置)

可以得到：

$$\mathbf{X_b} = \mathbf{R_{ba}} \mathbf{X_b} + \mathbf{t_{ba}}$$

带入 (1)式$\vec{X} = zK^{-1}\vec{p}$ 可以得到

$$z_b \mathbf{K}_b^{-1} \mathbf{p}_b = z_a \mathbf{R}_{ba} \mathbf{K}_a^{-1} \mathbf{p}_a + \mathbf{t}_{ba}$$

带入(*)式可以得到如下推导过程：

$$\begin{aligned} \mathbf{p}_b &= \frac{z_a}{z_b} \mathbf{K}_b \mathbf{R}_{ba} \mathbf{K}_a^{-1} \mathbf{p}_a + \frac{1}{z_b} \mathbf{K}_b \mathbf{t}_{ba} \\ &= \frac{z_a}{z_b} \mathbf{K}_b (\mathbf{R}_{ba} \mathbf{K}_a^{-1} \mathbf{p}_a + \frac{\mathbf{t}_{ba}}{z_a}) \\ &= \frac{z_a}{z_b} \mathbf{K}_b (\mathbf{R}_{ba} - \frac{\mathbf{t}_{ba} \vec{n}_a^T}{d_a}) \mathbf{K}_a^{-1} \mathbf{p}_a \end{aligned}$$

由于齐次坐标的尺度不变形 即可以省略$\lambda = \frac{z_a}{z_b}$，那么可以得到

$$\mathbf{p_b} \propto \mathbf{K}_b (\mathbf{R}_{ba} - \frac{\mathbf{t}_{ba} \vec{n}_a^T}{d_a}) \mathbf{K}_a^{-1} \mathbf{p}_a$$

所以$\mathbf{p_b} = \mathbf{H}_{ba} \mathbf{p}_a$ 可以得到 $\mathbf{H}_{ba} = \mathbf{K}_b (\mathbf{R}_{ba} - \frac{\mathbf{t}_{ba} \vec{n}_a^T}{d_a}) \mathbf{K}_a^{-1}$

由于变换矩阵 $\mathbf{T}_{ab} = \begin{pmatrix} \mathbf{R}_{ab} & \mathbf{t}_{ab} \\ \mathbf{0} & 1 \end{pmatrix}$,$\mathbf{T}_{ba} = \begin{pmatrix} \mathbf{R}_{ba} & \mathbf{t}_{ba} \\ \mathbf{0} & 1 \end{pmatrix}$ 并且 $\mathbf{T}_{ab} \mathbf{T}_{ba} = \mathbf{E}$

可以得到 $\mathbf{R}_{ba} \mathbf{t}_{ab} = -\mathbf{t}_{ba}$ 将其带入 $\mathbf{H}_{ba} = \mathbf{K}_b (\mathbf{R}_{ba} - \frac{\mathbf{t}_{ba} \vec{n}_a^T}{d_a}) \mathbf{K}_a^{-1}$

亦可以得到 $\mathbf{H}_{ba} = \mathbf{K}_b (\mathbf{R}_{ba} - \frac{\mathbf{t}_{ba} \vec{n}_a^T}{d_a}) \mathbf{K}_a^{-1}$

$$= \mathbf{K}_b (\mathbf{R}_{ba} + \frac{\mathbf{R}_{ba} \mathbf{t}_{ab} \vec{n}_a^T}{d_a}) \mathbf{K}_a^{-1}$$

$$= \mathbf{K}_b \mathbf{R}_{ba} (\mathbf{E} + \frac{\mathbf{t}_{ab} \vec{n}_a^T}{d_a}) \mathbf{K}_a^{-1}$$

获取方式

- 如果我们知道相机的内参数$\mathbf{K}_a, \mathbf{K}_b$和两个相机的旋转和平移变换矩阵$\mathbf{R}_{ab}, \mathbf{R}_{ba}, \mathbf{t}_{ab}, \mathbf{t}_{ba}$，我们就可以直接求出单应性变换矩阵$\mathbf{H}_{ab}$ 或者$\mathbf{H}_{ba}$
- 如果我们不知道各个参数矩阵那么可以采用4对匹配点法进行计算求解：

1. 由于 $p_{b1} \propto Hp_{a1}$

$$\begin{pmatrix} u_{b1} \\ v_{b1} \\ 1 \end{pmatrix} => \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} u_{a1} \\ v_{a1} \\ 1 \end{pmatrix}$$

2. 通过变换可以得到如下的形式：

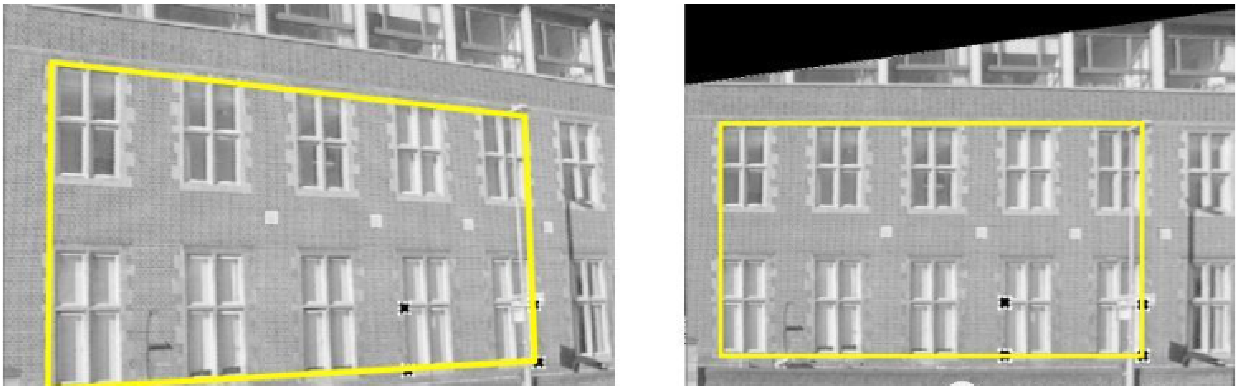
$$\begin{bmatrix} u_{a1} & v_{a1} & 1 & 0 & 0 & 0 & -u_{a1}u_{b1} & -v_{a1}u_{b1} & -u_{b1} = 0 \\ 0 & 0 & 0 & u_{a1} & v_{a1} & 1 & -u_{a1}v_{b1} & -v_{a1}v_{b1} & -v_{b1} = 0 \\ u_{a2} & v_{a2} & 1 & 0 & 0 & 0 & -u_{a2}u_{b2} & -v_{a2}u_{b2} & -u_{b2} = 0 \\ 0 & 0 & 0 & u_{a2} & v_{a2} & 1 & -u_{a2}v_{b2} & -v_{a2}v_{b2} & -v_{b2} = 0 \\ u_{a3} & v_{a3} & 1 & 0 & 0 & 0 & -u_{a3}u_{b3} & -v_{a3}u_{b3} & -u_{b3} = 0 \\ 0 & 0 & 0 & u_{a3} & v_{a3} & 1 & -u_{a3}v_{b3} & -v_{a3}v_{b3} & -v_{b3} = 0 \\ u_{a4} & v_{a4} & 1 & 0 & 0 & 0 & -u_{a4}u_{b4} & -v_{a4}u_{b4} & -u_{b4} = 0 \\ 0 & 0 & 0 & u_{a4} & v_{a4} & 1 & -u_{a4}v_{b4} & -v_{a4}v_{b4} & -v_{b4} = 0 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix}$$

因为齐次坐标，令$h_{33} = 1$，由此我们只需要求解H中的八个未知数，因此我们只需要找到4对匹配点，便可以计算出H矩阵的解，从而得到单应性矩阵H。

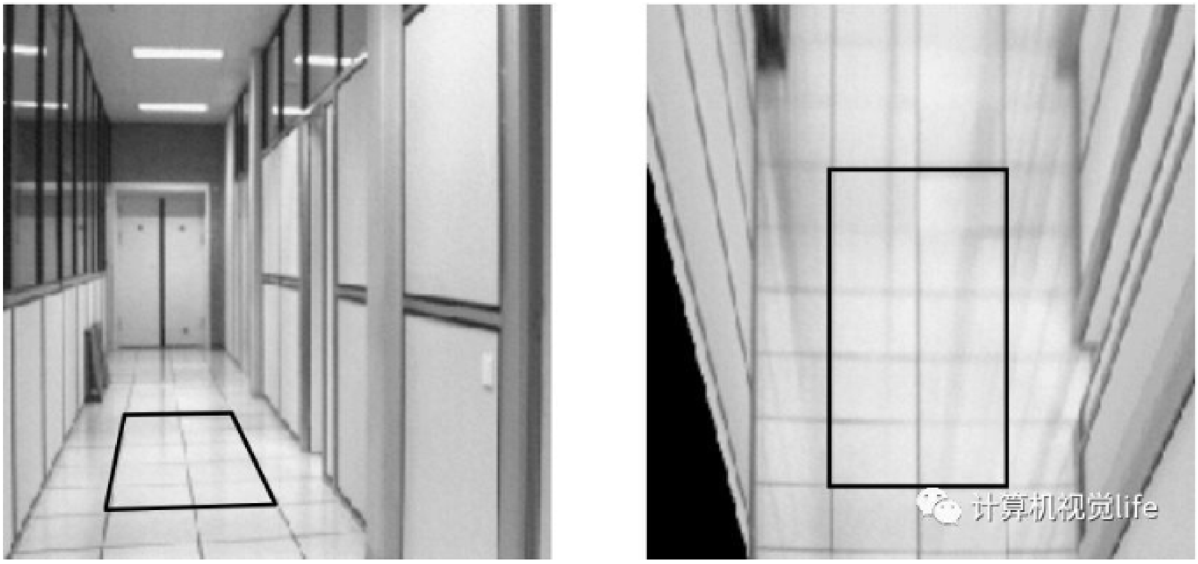
应用

单应性矩阵是两幅图像之间的变换关系，对此可知单应性矩阵在图像校正、图像拼接、相机位姿估计、视觉SLAM等领域有非常重要的作用

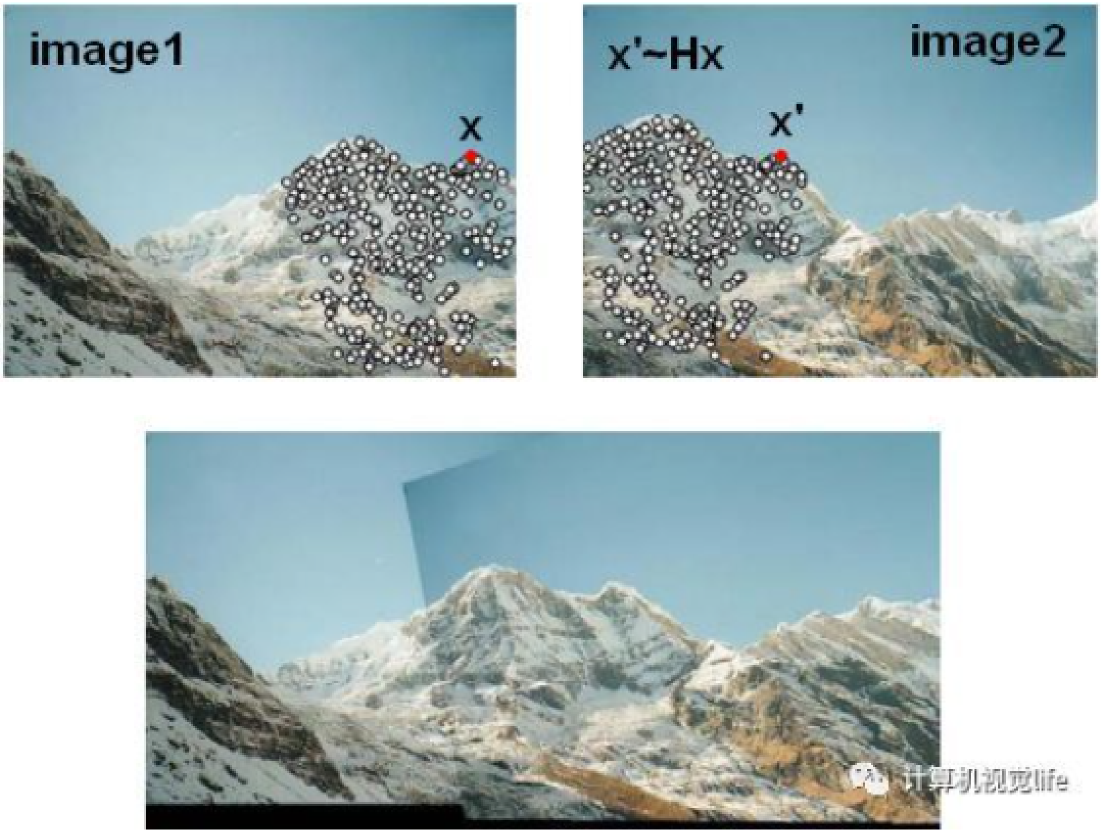
图像校正： 用单应性矩阵进行图像校正，如下图所示，至少需要四个对应点就可以实现。



视角变换： 单应性矩阵用于视角变换，如下图所示，可将左边普通视图转换为右边的俯瞰视图



图像拼接： 使用单应性矩阵，将不同角度拍摄的图像都转换到同一视角下，实现图像拼接



第3题

预知识

射影几何中2D点的表示方式 $\hat{p}=\begin{pmatrix}u\\v\\1\end{pmatrix}$,线的表示方式 $l=\begin{pmatrix}a\\b\\c\end{pmatrix}$,点在线上可以表示为

$$\hat{p}^T l=\begin{pmatrix}u & v & 1\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix}=au+bv+c=0$$

如果另一个点$\hat{p}'$也在线l上，那么$\hat{p}'^{T} l = \begin{pmatrix} u' & v' & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = au' + bv' + c = 0$

从而 $l = \hat{p} \times \hat{p}'$

相机坐标系下的坐标为 $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

推导过程

基础矩阵 $\mathbf{F}$:是通过对极几何一副图像上的**点**确定为另一幅图像上的**一条直线** 如 $l' = \mathbf{F}\hat{p}$

本质矩阵$\mathbf{E}$： 是基础矩阵 $\mathbf{F}$ 的一种特殊情况，是基础矩阵在归一化图像上的基础矩阵

反映空间中一点P在不同视角摄像机的**相机坐标系**下表示之间的关系 $X' = \mathbf{E}X$

由第二题和第三题的预习知识可知：

$p_1 = K_1 X_1$, 则可以得到 $X_1 = K_1^{-1} p_1$ ，

并且由刚体变换可知$X_2 = RX_1 + t$，所以$p_2 = K_2 X_2 = K_2(RX_1 + t) = K_2(RK_1^{-1}p_1 + t)$

对两边同时乘以K_2^{-1}可以得到 $K_2^{-1}p_2 = RK_1^{-1}p_1 + t$

设$x_1 = K_1^{-1}p_1, x_2 = K_2^{-1}p_2$ 带入到上式可以得到$x_2 = Rx_1 + t$

两边同时左叉乘t,由于txt =0

可以得到 $t \times x_2 = t \times Rx_1$

再同时左乘x_2^T 可以得到 $x_2^T t \times x_2 = x_2^T t \times Rx_1$

由于$t \times x_2$ 结果垂直与x_2 ,所以 $x_2^T \cdot (t \times x_2) = 0$,所以 $x_2^T t \times Rx_1 = 0$

将$x_1 = K_1^{-1}p_1, x_2 = K_2^{-1}p_2$ 带入上式可以得到 $p_2^T K_2^{-T} t \times RK_1^{-1} p_1 = 0$

所以$F = K_2^{-T} t \times RK_1^{-1}$**,并且** $\mathrm{E} = t \times R$

求法：

基础矩阵F -八点法

由于$l_2 = Fp_1$,并且 $p_2^T l_2 = p_2^T F p_1 = 0$，所以

$$[u_1, v_1, 1] \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = 0$$

经过处理之后

$$
\begin{bmatrix}
u_1^1 u_2^1 & u_1^1 v_2^1 & u_1^1 & v_1^1 u_2^1 & v_1^1 v_2^1 & v_1^1 & u_2^1 & v_2^1 & 1 \\
u_1^2 u_2^2 & u_1^2 v_2^2 & u_1^2 & v_1^2 u_2^2 & v_1^2 v_2^2 & v_1^2 & u_2^2 & v_2^2 & 1 \\
u_1^3 u_2^3 & u_1^3 v_2^3 & u_1^3 & v_1^3 u_2^1 & v_1^3 v_2^1 & v_1^3 & u_2^3 & v_2^3 & 1 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
u_1^8 u_2^8 & u_1^8 v_2^8 & u_1^8 & v_1^8 u_2^8 & v_1^8 v_2^8 & v_1^8 & u_2^8 & v_2^8 & 1
\end{bmatrix}
\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = 0
$$

从而可以计算出来基础矩阵F

本质矩阵--八点法

$$
\boldsymbol{x}_2^T E \boldsymbol{x}_1 = [u_2, v_2, 1] \begin{bmatrix} e_0 & e_1 & e_2 \\ e_3 & e_4 & e_5 \\ e_6 & e_7 & e_8 \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = 0
$$

经过带入和变换可以得到如下形式的方程组：

$$
\begin{bmatrix}
u_2^1 u_1^1 & u_2^1 v_1^1 & u_2^1 & v_2^1 u_1^1 & v_2^1 v_1^1 & v_2^1 & u_1^1 & v_1^1 & 1 \\
u_2^2 u_1^2 & u_2^2 v_1^2 & u_2^2 & v_2^2 u_1^2 & v_2^2 v_1^2 & v_2^2 & u_1^2 & v_1^2 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
u_2^n u_1^n & u_2^n v_1^n & u_2^n & v_2^n u_1^n & v_2^n v_1^n & v_2^n & u_1^n & v_1^n & 1
\end{bmatrix}
\begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \end{bmatrix} = 0
$$

从而可以求解出本质矩阵E

意义

基础矩阵F

可以将一个摄像机下的图像点与另一个摄像机图像上对应极线进行对应起来$l' = \mathbf{F}\hat{p}$

由上述推导过程可以知道 $p_2^T\mathbf{F}p_1 = 0$ 并且图像点p_1在第二幅上的对应点p_2,在极线$l_2 = \mathbf{F}p_1$上

基本矩阵F的自由度为7个，秩为2

基本矩阵F和相机内外参数关系为 $F = K_2^{-T}t \times RK_1^{-1}$，其中 K_1, K_2 为内参数矩阵，R,t为外参数矩阵

基本矩阵不依赖于摄像机矩阵的选择，等价地说不依赖于世界坐标系的选择

本质矩阵E

描述了两幅**规范化图像间**的极几何，**它与基本矩阵一样也是一个秩为 2,并且自由度为5 的矩阵**

与基本矩阵的不同之处是它仅与摄像机的运动参数有关。因此，从本质矩阵出发可估计出摄像机的欧氏运动参数

由上述推导过程可以知道基础矩阵F和本质矩阵E的关系为 $\mathbf{F} = K_2^{-T}\mathbf{E}K_1^{-1}$

第4题

相机内参数 K

为相机的内参矩阵，其中$\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$

内参矩阵反应了相机自身的属性，各个相机是一不一样的，需要标定才能知道这些参数

外参矩阵是[R|t]

反应了世界坐标系到相机坐标系的转换，与相机自身的属性无关。

联系

由2、3题目可以知道

$$\begin{aligned} \text{单应性变换矩阵 } \mathbf{H}_{ba} &= \mathbf{K}_b(\mathbf{R}_{ba} - \frac{\mathbf{t}_{ba}\vec{n}_a^T}{d_a})\mathbf{K}_a^{-1} \\ &= \mathbf{K}_b\mathbf{R}_{ba}(\mathbf{E} + \frac{\mathbf{t}_{ab}\vec{n}_a^T}{d_a})\mathbf{K}_a^{-1} \end{aligned}$$

基础矩阵$F = K_2^{-T}t \times RK_1^{-1}$

本质矩阵 $\mathbf{E} = t \times R$

可以知道 H，F矩阵和相机的参数矩阵(内外参数K，R,t都有关系)，而本质矩阵与相机的内参数无关，只和相机的外参数有关系。

区别

相机的参数矩阵反映了相机的内在属性(内参矩阵)，以及世界坐标和相机坐标之间的转换关系(外参矩阵)

单应性变换矩阵H反应的是两幅图像**像素坐标**之间的关系，可以直接根据单应矩阵H从而计算出一副图像上的点在另一副图像上点的像素坐标，被广泛应用在图像处理的各个方面

基础矩阵F可以将一个摄像机下的图像点与另一个摄像机图像上对应极线进行对应起来

本质矩阵E是基础矩阵的一个特殊形式，反应的是**图像坐标**之间直接对应关系

第5题

Rectification：整流也称为立体校正

为什么

因为当两个图像平面是完全共面行对准时，**计算立体视差是最简单的**。但是，在现实的双目立体视觉系统中，是**不存在完全的共面行对准的两个摄像机图像平面的**，所以我们要进行立体校正

图像立体校正是指对两幅图像分别进行一次平面投影变换，使两幅图像的对应极线在同一条水平向上，而对极点被映射到无穷远处， 这样可以使两幅图像只存在水平方向上的视差，立体匹配问题从二维降到一维，**从而提高了匹配的速度**、

步骤：

输入：K1, R1, t1, K2, R2, t2

1. 求光心位置：

$$c_1 = -R_1^T K_1^{-1} \bar{t_1}$$
$$c_2 = -R_2^T K_2^{-1} \bar{t_2}$$

2. 确定新相机旋转矩阵

$$r_1 = \frac{c_1 - c_2}{||c_1 - c_2||}$$
$$r_2 = R_1(3,:) \times r_1$$
$$r_3 = r_1 \times r_2$$
$$R_n = \begin{bmatrix} r_1^T \\ r_2^T \\ r_3^T \end{bmatrix}$$

3. 求新相机内参矩阵

$$K_n = \frac{K_1 + K_2}{2}$$

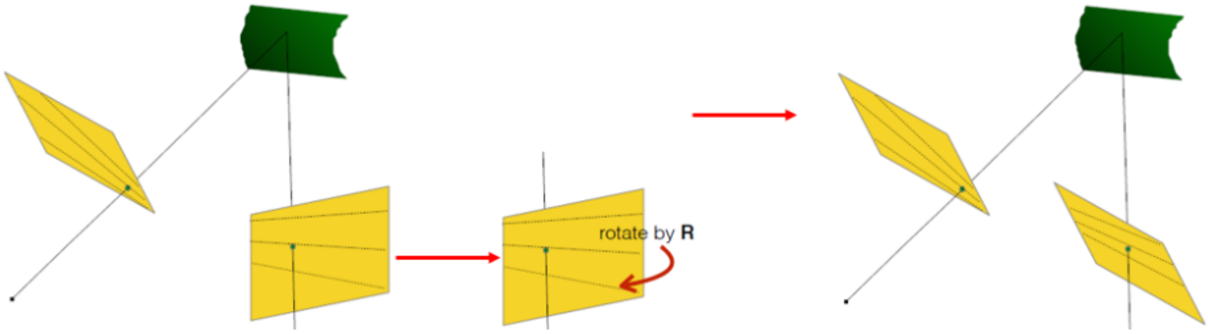
4. 求变换矩阵

$$T_1 = K_n R_n R_1^T K_1^{-1}$$
$$T_2 = K_n R_n R_2^T K_2^{-1}$$

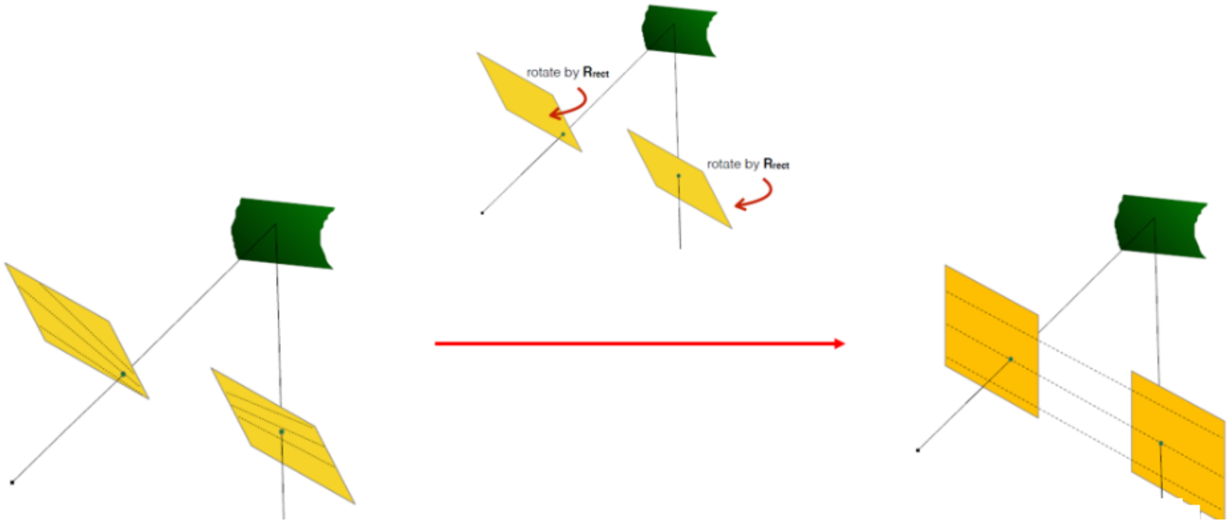
5. 用变换矩阵T1和T2对两个图像进行变换

在变换的过程中分为主要分为两个步骤：

1、第一次旋转：对右图进行旋转，从而使得右图像平面和左图像平面平行



2、第二次旋转:通过旋转从而建立一个新的坐标系，使得新的像平面共面并且平行于相机平面



第6题

SFM： 是从一系列包含视觉运动信息的多幅二维图像序列中估计三维结构的技术，可以由图像生成3维点云以及相机姿态，SFM是三维重建中的一项重要技术。
- 输入：一系列图像，拍摄同一场景
- 输出： 每一张图像对应的相机位置和朝向， 场景中的3D点云

算法过程

1. 相机标定，获取相机内参
2. 使用标定好的相机从多个角度拍摄同一场景图片，并按序号进行保存；
3. 对相邻图像两两计算匹配特征点；一般首先使用两张图像进行重建，计算出一个初始的点云，之后不断添加后续的图像，具体添加哪一张图像的方法是：检查已有的图像中哪一个与已有点云中的点匹配最多，就选哪一张；
4. 使用3中计算好的对应点对计算基础矩阵F
5. 通过基础矩阵计算本质矩阵E;
6. 通过本征矩阵计算两个视角之间的运动，即R，T
7. 在计算出[R|T]矩阵后，就可以使用光学三角法对所有的特征点重建

应用实例

SFM与摄影测量

SFM中采用SIFT等特征提取算法进行特征点的提取与匹配是两者之间共有的，SFM中的极线约束在摄影测量里基于核线的影像匹配的思路一致，并且SFM中的三角化对应于摄影测量中的前方交会，SFM中的Bundle Adjustment正是摄影测量中的光束法平差的思路，因此SFM可以充分应用在摄影测量的方面，如航空摄影测量。