Q1. A homography H can be multiplied by a constant without changing the mapping as it maps homogeneous points to homogeneous points. So, homographies have 8 degrees of freedom. We might be tempted to say that one of the elements of H has a particular value, say H_{33}= 1. However, this is not always possible, for example, there are homographies for which H_{33}= 0. What can you say in general about a homography H for which H_{33}= 0? (Hint: Think about the which point in the scene will be mapped to origin in the image)
Sol.

$$H \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix} = \begin{vmatrix} H_{13} \\ H_{23} \\ 0 \end{vmatrix} \tag{1}$$

So, the origin is mapped to a point at infinity. An example is the following. Suppose the first plane is an image projection plane and the second plane is a scene plane, so that H(inverse) maps from the image plane coordinates to the scene plane coordinates. The mapping takes the origin of the image plane to a point at infinity on the scene plane. This means that the origin in the image plane lies on the vanishing line of the scene plane i.e. the origin lies on the horizon

Q2. Please write down what the explicit values of the parameters defining the internals of a camera: focal length, the principal point, skew, and scaling in the x, y-directions. The internal matrix for the camera is

$$K1 = \begin{pmatrix} 10 & 0 & 10 \\ 0 & 10 & 10 \\ 0 & 0 & 1 \end{pmatrix}$$

Sol: $f = 10$; skew is zero; the principal point is $(10, 10)$; scaling in the x and y directions is 1

Q3. An ideal pinhole camera has focal length 5mm. Each pixel is 0.02 mm $\times 0.02$ mm and the image principal point is at pixel $(500, 500)$. Pixel coordinates start at $(0, 0)$ in the upper-left corner of the image. What is the 3×3 camera calibration matrix, K, for this camera configuration?
Sol:

$$K1 = \begin{pmatrix} 5 & 0 & 500 * 0.02 \\ 0 & 5 & 500 * 0.02 \\ 0 & 0 & 1 \end{pmatrix}$$

Q4. You are trying to fit a homography between two images. Unfortunately, due to the nature of images, SIFT returned a number of false correspondences which amounts to 40% of the total correspondences. How many iterations should you run the RANSAC algorithm so that the probability of getting the correct homography is larger than 0.99?

Sol: As long as all the points RANSAC picks in a particular iteration are inliers, it will the fit the right homography. The probability that it will pick all inliers in an iteration is 0.6^4. Over n iterations, the probability that it will pick all inliers in at least one iteration is $1-(1-0.6^4)^n$. So the required n is,

$$\begin{aligned} 1-(1-0.6^4)^n &> 0.99 \\ 0.01 &> (1-0.6^4)^n \\ -2 &> n\log(1-0.6^4) \\ n &> \frac{-2}{\log(1-0.6^4)} \\ &\approx 33.18 \end{aligned}$$

So, RANSAC has to be run for at least 34 iterations which is surprisingly low!

Q5.

a. **A scene point at coordinates (400,600,1200) is perspectively projected into an image at coordinates (24,36), where both coordinates are given in millimeters in the camera coordinate frame and the camera's principal point is at coordinates (0,0,f) (i.e., $u_0 = 0$ and $v_0 = 0$). Assuming the aspect ratio of the pixels in the camera is 1, what is the focal length of the camera? (Note: the aspect ratio is defined as the ratio between the width and the height of a pixel; i.e., ku/kv.)**

b. **Show how the projection of a point in a planar scene at world coordinates (X, Y) to pixel coordinates (u, v) in an image plane can be represented using a planar affine camera model.**

Sol:

$$\begin{aligned} u &= \frac{fx}{z} \\ f &= \frac{uz}{x} = \frac{24 \times 1200}{400} = 72 \text{ mm} \end{aligned}$$

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ 0 & 0 & p_{33} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

(The overall scale does not matter)

Q6. Let A be a real skew-symmetric matrix, that is, $A^T = -A$. Then prove the following statements.

a **Each eigenvalue of the real skew-symmetric matrix A is either 0 or a purely imaginary number.**

b **The rank of A is even.**

Sol: (a) Let λ be an eigenvalue of A and x be an eigenvector corresponding to the eigenvalue λ. That is, we have $Ax = \lambda x$.

Multiplying by $\hat{x}^T$ from the left, we have $\hat{x}Ax = \lambda\hat{x}^T x = \lambda||x||^2$.

Since dot product is commutative and A is skew-symmetric ($A^T = -A$), we have

$$\hat{x}Ax = (Ax)^T\hat{x} = A^T x^T \hat{x} = -x^T A\hat{x}$$

.

Taking cojugate of $Ax = \lambda x$ and use the fact that A is real, we have $A\hat{x} = \hat{\lambda}\hat{x}$.

Thus, we have

$$-x^T A\hat{x} = -x^T\hat{\lambda}\hat{x} = -\hat{\lambda}||x||^2$$

.

Therefore, $-\hat{\lambda}||x||^2 = \lambda||x||^2$.

Since x is an eigenvector, it is nonzero by definition. Hence we have $-la\hat{m}bda = \lambda \implies \lambda$ is either 0 or purely imaginary

(b) From part(a), we know that the eigenvalues of A are 0 are purely imaginary.

Thus if λ is a purely imaginary eigenvalue, then its conjugate $\hat{\lambda} = -\lambda$ is also an eigen value of A since A is real matrix. Thus, nonzero eigenvalues come in pairs $\lambda, -\lambda$.

Let $\lambda_1, -\lambda_1 \ldots, \lambda_k, -\lambda_k$ be nonzero eigenvalues of A.

Since a real skew-symmetric matrix is normal, it is diagonalizable (by a unitary matrix). Thus there exist an invertible matrix P such that $P^{-1}AP = diag[\lambda_1 \ -\lambda_1 \ \ldots \ \lambda_k \ -\lambda_k \ 0 \ \ldots \ 0]$.

Since P is invertible matrix, rank of A is same as rank of the diagonal matrix on the right-hand side which is $2k$.

Q7.

a. **what happens if the aperture size is expanded/shrinked in the pin-hole camera model?**

b. **what is the relation between aperture size and depth of field?**

Sol: (a)If the aperture is expanded, more amount of light rays are passed which leads to blurring. If the aperture is shrinked more, less light gets through resulting in noisy and diffraction effects.

(b) A smaller aperture increases the range in which the object is approximately in focus

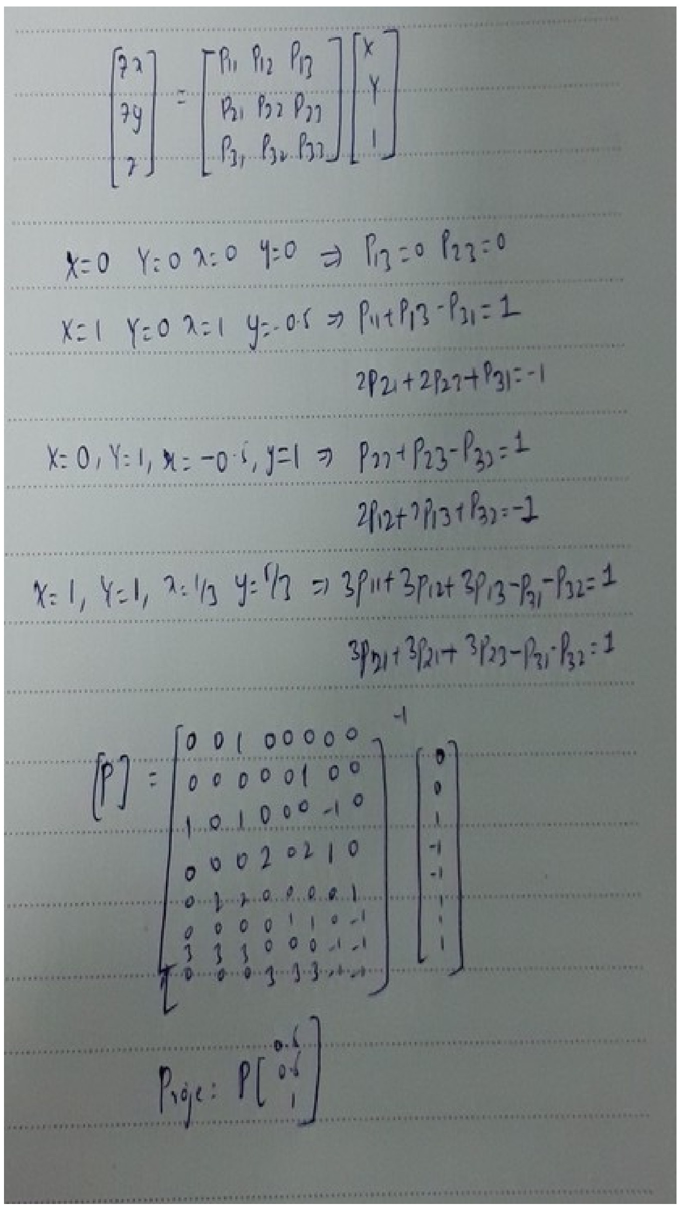
Q8. Given that a unit square in the world plane

$$X^p = \{(0,0),(1,0),(0,1),(1,1)\}$$

is mapped by perspective projection onto the image plane at

$$x = \{(0,0),(1,0.5),(0.5,1),(1/3,1/3)\}$$

find the planar projective transformation that describes the mapping from X^p to x. What is the image of the point Xp = (0.5, 0.5)?



$$\begin{bmatrix} zx \\ zy \\ z \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

$X=0 \quad Y=0 \quad x=0 \quad y=0 \Rightarrow P_{13}=0 \quad P_{23}=0$

$X=1 \quad Y=0 \quad x=1 \quad y=-0.5 \Rightarrow P_{11}+P_{13}-P_{31}=1$

$2P_{21}+2P_{23}+P_{31}=-1$

$X=0,\ Y=1,\ x=-0.5,\ y=1 \Rightarrow P_{22}+P_{23}-P_{32}=1$

$2P_{12}+2P_{13}+P_{32}=-1$

$X=1,\ Y=1,\ x=1/3 \quad y=1/3 \Rightarrow 3P_{11}+3P_{12}+3P_{13}-P_{31}-P_{32}=1$

$3P_{21}+3P_{22}+3P_{23}-P_{31}-P_{32}=1$

$$[P] = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 & 0 & 2 & 1 & 0 \\ 0 & 2 & 2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 \\ 3 & 3 & 3 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 3 & 3 & 3 & -1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\text{Proje}: P\begin{bmatrix} 0.5 \\ 0.5 \\ 1 \end{bmatrix}$$

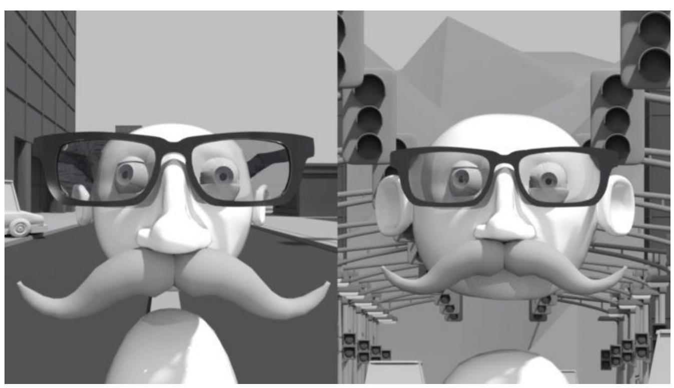
Q9. Say you have the following blurred (5×5) natural images

$$I_1 = \begin{bmatrix} 30 & 29 & 30 & 29 & 28 \\ 50 & 51 & 50 & 51 & 52 \\ 45 & 44 & 45 & 44 & 45 \\ 15 & 16 & 20 & 15 & 15 \\ 100 & 101 & 105 & 101 & 100 \end{bmatrix}, I_2 = \begin{bmatrix} 99 & 98 & 101 & 100 & 101 \\ 99 & 98 & 99 & 101 & 100 \\ 99 & 98 & 100 & 98 & 99 \\ 99 & 100 & 98 & 99 & 98 \\ 100 & 101 & 100 & 100 & 100 \end{bmatrix}$$

Which one of them has motion blur and which one has defocus blur and for the one with motion blur, in which direction is the motion happening?

Ans. In case of defocus blur, the gradient would be same in all the directions but for motion blur, the gradient along the direction of motion would be smooth whereas the ones perpendicular to the motion would be much sharper. So, I_2 has defocus blur and I_1 has motion blur. In I_1, the motion is horizontal.

Q10. Given the fact the angle of view (α) changes with focal length (f) according to the relation $\alpha = 2 \arctan \frac{d}{2f}$, which one of the following images was taken with an increased focal length and which one was taken by moving the camera closer?



Ans: Clearly, in the image to the right, the field of view is much more restricted than in the one to the left. So, the image to the right was taken by increasing the focal length of the camera while the other one was taken by moving the camera closer to the subject.

Q11. If I double the focal length of a camera lens and also move twice as far away from an object I focus on in a scene, what are two (2) things that will be different in the two images?

Sol: The fields of view are different (because the sensor size remains the same in both cases) and also the depths of field are different (causing different amounts of blur in the areas in front of or behind the object in focus).

Q12. Let $u = [u_1, u_2, 1]^T$ and $v = [v_1, v_2, 1]^T$ denote the homogeneous coordinates of points

in planes P and Q, respectively. Matching point pairs are related by a projection matrix H as:

$$v_i = Hu_i, \quad i = 1, 2, ..., n \tag{2}$$

A transform from P to Q in the form of a translation by (x_0, x_1) **then a rotation by** θ **then a scaling by** (s_0, s_1) **is described by the matrix:**

$$H = \begin{vmatrix} 0.95162 & 0.443749 & -6.97686 \\ -0.40148 & 0.860992 & -2.29753 \\ 0 & 0 & 1 \end{vmatrix} \tag{3}$$

Determine the values of $(x_0, x_1, s_0, s_1, \theta)$**.**

$$\mathbf{H} = \mathbf{SRT} = \begin{bmatrix} s_0 & 0 & 0 \\ 0 & s_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & x_1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} s_0\cos\theta & s_0\sin\theta & s_0x_0\cos\theta + s_0x_1\sin\theta \\ -s_1\sin\theta & s_1\cos\theta & s_1x_1\cos\theta - s_1x_0\sin\theta \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore

$$\begin{aligned}
s_0 &= \left(H_{1,1}^2 + H_{1,2}^2\right)^{1/2} = \left(0.951623^2 + 0.401487^2\right)^{1/2} = 1.05 \\
s_1 &= \left(H_{2,1}^2 + H_{2,2}^2\right)^{1/2} = \left(0.401487^2 + 0.860992^2\right)^{1/2} = 0.95 \\
x_0 &= H_{1,3}\cos\theta/s_0 - H_{2,3}\sin\theta/s_1 = -5 \\
x_1 &= H_{1,3}\sin\theta/s_0 + H_{2,3}\cos\theta/s_1 = -5 \\
\theta &= \arctan(H_{1,2}, H_{1,1}) * 180/\pi = 25^\circ
\end{aligned}$$