Questions

1. **lecture 10: 2D edge detection, 2nd moment matrix**

   Give an example of a second moment matrix that has rank 1 and whose diagonal elements are different from each other and non-zero. Describe how local image intensities can produce such a matrix.

2. **lecture 11: Harris corners, image registration**

   (a) We discussed how to register two 1D images $I(x)$ and $J(x)$ by using a first order Taylor series expansion of $I(x+h)$. Express the solution in terms of a least squares problem of the form: "find the $\mathbf{x}$ that minimizes $\| \mathbf{Ax} - \mathbf{b} \|$."

   (b) Do the same for the 2D image registration problem discussed in the lecture 10 which was based on image translation.

   (c) A more general transformation that is commonly used to register two images is

   $$I(\mathbf{x} + \mathbf{D}(\mathbf{x} - \mathbf{x_0}) + \mathbf{h}) = J(\mathbf{x}).$$

   where $\mathbf{D}$ is a 2×2 image (called the "deformation"). Write this transformation as a homography.

   (d) Generalize the transformation of (c) so that it allows for any homography.

   (e) What are the advantages/disadvantages of using a more general model ?

3. **lectures 12-14: scale space**

   (a) What assumptions about the image intensity and image "motion" are made by the Lucas-Kanade image registration approach?

   (b) How does a multiscale approach (coarse-to-fine) allow us to weaken these assumptions?

   (c) One of the main conclusions of the 1D Canny edge analysis of lecture 8 is that using a bigger filter increases the signal to noise ratio (detection), but also increases the variance of position where the edge is found (localization). Given this observation, how might you improve edge detection by using multiscale approach ?

4. **lecture 15 - finding lines and vanishing points, Hough, RANSAC**

   (a) We discussed how to fit a line to a set of points in a plane using a Hough transform. What if we wish to fit lines using edge elements (x, y, θ) rather than just x, y, e.g. 2D Canny gives us such edge elements. How would the Hough method change?

   (b) Let two lines l_i be

   $$a_i x + b_i y + c_i = 0$$

   where $i = 1, 2$. In introductory linear algebra, you learn how to compute the intersection of the two lines using Gaussian elimination. There is also a quicker way, namely take the cross product, $(a_1, b_1, c_1) \times (a_1, b_1, c_1)$ and treat the result as a homogenous vector (so to get the intersection point, you divide by the 3rd coordinate). Give a geometric argument that explains this fact.

(c) Suppose we have the two lines

$$3x + 4y + 2 = 0$$

$$2x - y = 0.$$

Compute their intersection using the method of (b).

5. **lecture 16 - shading on a sunny day**

   Suppose a surface of height $Z_0 + Z(X,Y)$ is illuminated by a parallel source in direction (l_X, l_Y, l_Z). Show that, for an "inverted surface" of height $Z_0 - Z(X,Y)$, it is possible to choose a different light source direction which gives the same surface irradiance as the original surface. This is known as the *depth reversal ambiguity*.

   See `http://www.youtube.com/watch?v=iR9WVhiaIeY` for a fascinating example.

6. **lecture 18 - least squares and SVD**

   In lecture 18, we examined how to use the pseudoinverse of a matrix to solve a least squares problem of finding the n-vector $\mathbf{x}$ that minimizes $\| \mathbf{Ax} - \mathbf{b} \|$, where $\mathbf{A}$ is a given $m \times n$ and $\mathbf{b}$ is an m-vector.

   Replace $\mathbf{A}$ with its SVD $\mathbf{U\Sigma V}^T$ in the above problem definition and left multiply by $\mathbf{U}^T$ (which does not change the L_2 norm). Write out the solution for the two cases $\mathbf{b} = 0$ and $\mathbf{b} \neq 0$, using this new formulation.

7. **lecture 19 - Homographies**

   (a) Since a homography $\mathbf{H}$ maps homogeneous points to homogeneous points, you can muliply a homography by a constant without changing the mapping that it represents. This implies, in particular, that homographies have 8 degrees of freedom. You might be tempted to represent these 8 degrees of freedom explicitly by insisting that one of the elements of $\mathbf{H}$ 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 $\mathbf{H}$ for which $H_{33} = 0$?

   (b) Consider a plane:

   $$Z = Z_0 - Y \tan\theta$$

   which we get by rotating the plane $Z = 0$ by θ degrees about the X axis and then translating by $(0, 0, Z_0)$. We parameterize points on the rotated plane by (s,t) such that the origin $(s,t) = (0,0)$ is mapped to the 3D point $(x,y,z) = (0,0,Z_0)$, and the s axis is in the direction of the X axis.

   What is the homography $\mathbf{H}$ mapping $(s,t,1)$ to the image projection plane $Z = f$, under perspective projection?

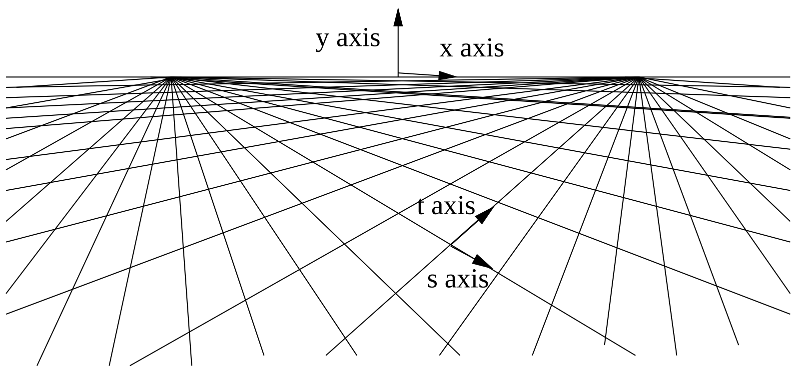
   (c) For (b), what is the vanishing line in the image?

   (d) Suppose the scene consists of a "ground plane", which in camera coordinates is

   $$Y = h \ < 0.$$

On the ground plane, let's assume we have a regularly spaced grid of square tiles, i.e. tiles on a floor. These define two sets of parallel lines. These parallel lines need not be aligned with the camera's **X** and **Z** unit vectors, but rather can be rotated by some angle θ relative to these unit vectors. Let the origin of the ground plane coordinate system be some point (X_0, h, Z_0).

What is the homography taking ground plane coordinates $(s, t, 1)$ to image plane coordinates, assuming a projection plane $Z = f$? What are the vanishing points that correspond to letting $s \to \infty$ (for finite t) and $t \to \infty$ (for finite s) ?



8. **lecture 22 - essential and fundamental matrix**

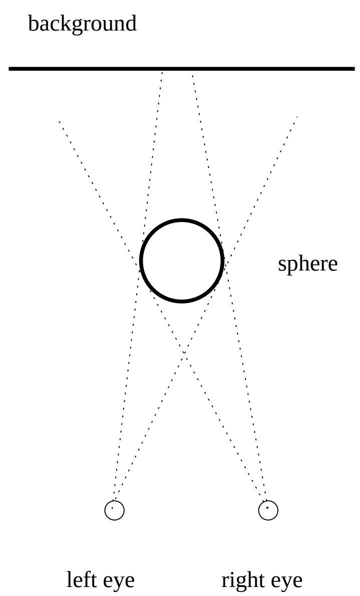
   (a) Suppose the first and second camera axes have the same directions and the second camera is located position $(T_X, 0, 0)$ in the first camera's coordinate system. What is the essential matrix **E** relating points in the two cameras' coordinate systems?

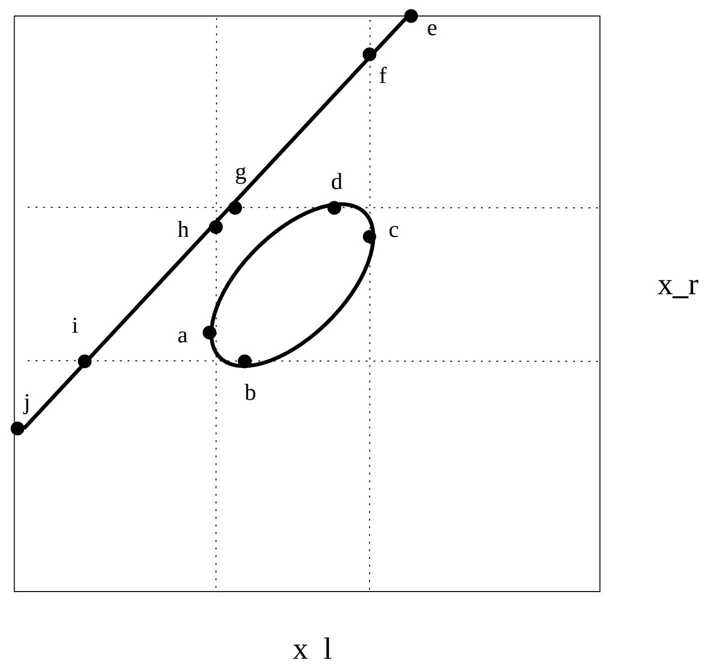
   (b) Show that, in general, an essential matrix has two non-zero singular values and they are identical. How does their value depend on the **R** and **T** matrices that relate the two cameras ?

   (c) Given a fundamental matrix **F** that relates corresponding points in two images, how are the epipoles related to the elements of the **SVD** of **F** ?

9. **lecture 23: stereo correspondence**

   Consider a sphere in front of a background. The figure below shows a slice through this scene, lying in an epipolar plane. Also shown is an (x_l, x_r) representation of points in the scene. Indicate for each surface region whether it is seen by both eyes, one eye (which one?) or neither eye.


background
sphere
left eye
right eye


e
f
g
d
h
c
i
a
j
b
x_r
x_l

Answers

1. **lecture 10: 2nd moment matrix**

   An example can be obtained by taking the outer product of any vector with itself e.g.

$$\begin{bmatrix} 5 \\ -2 \end{bmatrix} \begin{bmatrix} 5 & -2 \end{bmatrix} = \begin{bmatrix} 25 & -10 \\ -10 & 4 \end{bmatrix}$$

   i.e. you would get such a second moment matrix if the intensity gradient vector was constant (5,-2) over the neighborhood.

2. **lecture 11: image registration**

   (a) In the 1D case we had

$$\sum_{x \in Ngd(x_0)} (I(x) - J(x) + h\frac{dI(x)}{dx} \ )^2.$$

   Let the neighborhood have N pixels, then we have an $N \times 1$ matrix $\mathbf{A}$ of $\frac{dI(x)}{dx}$ values, such that $\mathbf{A}^T\mathbf{A} = \sum_{x \in Ngd(x_0)} (\frac{dI(x)}{dx} \ )^2$. (Here we are ignoring the weighting function for the neighborhood.) The $\mathbf{x}$ "vector" is in this example just a single variable h, i.e. a scalar, and the $\mathbf{b}$ vector is an $N \times 1$ vector of $(I(x) - J(x))$ values.

   (b) For the 2D case, we are minimizing

$$\sum_{(x,y) \in Ngd(x_0,y_0)} (I(x,y) - J(x,y) + \frac{\partial I}{\partial x}h_x + \frac{\partial I}{\partial y}h_y)^2 = \parallel \mathbf{A}\mathbf{x} - \mathbf{b} \parallel^2$$

   where the neighborhood has N pixels, $\mathbf{A}$ is an $N \times 2$ matrix of $(\frac{\partial I(x,y)}{\partial x} \ , \frac{\partial I(x,y)}{\partial y})$ values, the $\mathbf{x}$ vector is (h_x, h_y), and the $\mathbf{b}$ vector is an $N \times 1$ vector $(J(x,y) - I(x,y))$. In particular, we wish to solve $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$.

   (c)

$$I(\mathbf{x} + \mathbf{D}(\mathbf{x} - \mathbf{x_0}) + \mathbf{h}) = J(\mathbf{x}).$$

   The transformation is therefore

$$\mathbf{x} \rightarrow \mathbf{x} + \mathbf{D}(\mathbf{x} - \mathbf{x_0}) + \mathbf{h}$$

   where $\mathbf{x} = (x, y)$. This transformation can be written in homogenous coordinates as

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 + D_{11} & D_{12} & h_x \\ D_{21} & 1 + D_{22} & h_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

   where the 3×3 matrix is invertible, hence it is a homography.

(d) The transformation referred to in the previous question is not as general as it could be since the first and second elements of the third row are 0, and the third element is 1. To make the tranformation model general enough to include all homographies, we could use

$$I(\mathbf{Hx}) = J(\mathbf{x})$$

where the $\mathbf{x}$ is written in homogeneous coordinates, and we could then try to find the $\mathbf{H}$ that minimizes the sum of squared errors over $\mathbf{x}$ in some local neighborhood of $\mathbf{x}_0$. (Further details omitted. If you are interested, see the excellent paper "Lucas-Kanade 20 years on: a unifying framework" by S. Baker and I. Matthews IJCV 2004.)

(e) The advantage of using a more general model is that it allows you to obtain a better fit to the data. The more restricted model fixes the value of some of the elements in the homography, in particular, the third row. By allowing these elements to vary when we find a least squares fit, we are guarenteed to get a better fit (or no worse, anyhow).

Another advantage is that by using a more general model, we sometimes allow ourselves to use a larger neighborhood. For example, a pure 2D translation model might be fine for a very small neighborhood, but if there is a stretching or shearing in addition to translation then pure 2D translation would not capture this. Similarly, if the true deformation involves perspective viewing of a plane, then you would likely need the full homography to capture this deformation.

What are the disadvantages in using a more general model? First, using more parameters causes the computation to be more expensive – this can be significant when we perform the computation at each pixel. Second, the more general model might be inappropriate in some situations, in which case the extra parameters might just "fit the noise". For example, if the deformation is caused by a non-planar object viewed by two different cameras (hence there is a 3D translation between the camera positions), then the transformation might not be well explained by a homography.

Keep in mind that "goodness of fit" just mean getting a smaller data error (i.e. image intensity error). This is not necessarily the ultimate goal. We might actually want to know the geometric transformation, since it might tell us something about the scene geometry.

3. (a) Lucas-Kanade assumes:

i. images have been smoothed so that we can reliably estimate the intensity gradient, in particular, the image gradient ∇I is locally constant over the motion distance $|(h_x, h_y)|$ so that the Taylor series approximation holds;

ii. the direction of the image intensity gradient varies over a σ_I neighbhorhood, so the second moment matrix is invertible;

iii. the motion vector (h_x, h_y) is constant over the σ_I neighborhood (i.e. we are solving for a single vector (h_x, h_y) for that neighborhood),

Note that for all these assumptions to hold, the magnitude of the motion vector (h_x, h_y) must be much smaller than the width of the σ_I neighborhood, since otherwise the assumptions would be mutually contradictory.

(b) If the motion (h_x, h_y) has a large magnitude, then we need to use a large blurring radius σ_D to satisfy assumption (i). But if we use a large σ_D, then we need to use a large σ_I to satisfy assumption (ii). In turn, if σ_I is large, then it is less likely that (iii) will hold. This seems to create a situation that is difficult to satisfy.

The multi-scale approach attempts to avoid this problem. It does so by first estimating *roughly* the translation (h_x, h_y) for large scales (large σ_D and σ_I), in particular, it assumes a constant single translation over the large neighborhood. It then shifts this image by this (approximate) translation, and attempts to re-estimate the (residual) translation for the shifted image, and using a smaller scale. Why should this work? If the large scale estimate of (h_x, h_y) is close to accurate, then the images can be approximately registered, such that the residual motion required to fully register them is of a very small magnitude.

(c) Briefly: Use a large σ (and hence large S:N) to detect edges. Then, reduce σ to better localize these detected edges, i.e. keep track of the edges that you've detected as you shrink σ.

Reducing the scale σ will lower the S:N, and you will detect more peaks the first derivative magnitudes (or gradient magnitudes in the 2D case), but the location of the edges that were detected edges at the larger scale will be more accurate. So, for each edge found at the larger scale, find the closest peak at the smaller scale, and ignore other peaks. If the scales σ are sampled finely, then you will be able to track the peaks from coarse to fine scales, obtaining excellent localization. (That is the idea, anyhow.)

4. (a) The Hough method used a `for` loop to consider all theta values. If we have an estimate of theta, then we would cast votes only for that theta (or a small range near that theta) rather than for all theta.

(b) Consider the 3D vectors (a_i, b_i, c_i) and $(x, y, 1)$, where the latter lies on a projection plane $Z = 1$. Because their dot product is 0, vectors (a_i, b_i, c_i) and $(x, y, 1)$ are orthogonal to each other. In particular, (a_i, b_i, c_i) is orthogonal to the plane π_i spanned by the origin and the line l_i. The two lines $l_i, i = 1, 2$ thus define two planes π_1 and π_2, both of which pass through the origin. The intersection of these two planes must therefore be a line that passes through the origin i.e. a "line of sight". This line meets $Z = 1$ at precisely the intersection of l_1 and l_2.

Since (a_i, b_i, c_i) is orthogonal to plane π_i, it follows that the cross product vector

$$(a_1, b_1, c_1) \times (a_2, b_2, c_2)$$

must lie in *both planes* π, and hence it lies on the intersection of the two planes. Thus, to get the intersection of the two lines l_1 and l_2, we compute their cross product (x, y, z) and normalize it so that it intersects $Z = 1$, and so the intersection point is $(\frac{x}{z}, \frac{y}{z})$

(c) Then $(3, 4, 2) \times (2, -1, 0) = (2, 4, -11)$ and so the intersection point is $(-\frac{2}{11}, -\frac{4}{11})$.

5. Inverting the depth inverts the depth gradient $(\frac{\partial Z}{\partial X}, \frac{\partial Z}{\partial Y}) \to -(\frac{\partial Z}{\partial X}, \frac{\partial Z}{\partial Y})$. Instead of light source direction (l_X, l_Y, l_Z), consider light source direction $(-l_X, -l_Y, l_Z)$, and note

$$(\frac{\partial Z}{\partial X}, \frac{\partial Z}{\partial Y}, -1) \cdot (l_X, l_Y, l_Z) = (-\frac{\partial Z}{\partial X}, -\frac{\partial Z}{\partial Y}, -1) \cdot (-l_X, -l_Y, l_Z)$$

Hence inverting the depth and inverting the XY coordinates of the light source gives the same surface irradiance.

6. We are now minimizing $\| \Sigma \mathbf{V}^T \mathbf{x} - \mathbf{U}^T \mathbf{b} \|$.

   The first version is that $\mathbf{b} \neq \mathbf{0}$, i.e. at least one element of the vector $\mathbf{b}$ is non-zero. Set $\mathbf{x}' = \mathbf{V}^T \mathbf{x}$ and $\mathbf{b}' = \mathbf{U}^T \mathbf{b}$. Note that at least one component of $\mathbf{b}'$ must be non-zero since $\mathbf{b}$ is non-zero and the columns of $\mathbf{U}^T$ are linearly independent.

   Let x_i', b_i' be the elements of $\mathbf{x}$' and $\mathbf{b}$', respectively. We want to find the x_i that minimize $\sum_{i=1}^{n} (\sigma_i x_i' - b_i')^2$. The solution is $x_i' = \frac{b_i'}{\sigma_i}$ if $\sigma_i > 0$ and $x_i' = 0$ if $\sigma_i = 0$.

   In the second version, $\mathbf{b} = \mathbf{0}$. In this case, we are trying to minimize $\| \Sigma \mathbf{V}^T \mathbf{x} \|$. It is trivial to minimize this by setting $\mathbf{x} = 0$, but this is not an interesting case. So we perform the minimization but require that $\| \mathbf{x} \| = 1$. By inspection, we get the minimum when $\mathbf{x}' = \mathbf{V}^T \mathbf{x}$ is the unit vector corresponding to the smallest singular value, so $x_i' = 0$ for all $i \neq n$ and $x_n' = 1$, i.e. so $\| \Sigma \mathbf{V}^T \mathbf{x} \| = \sigma_n$.

7. (a) If $H_{33} = 0$, then in particular we would have

$$\mathbf{H} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} H_{13} \\ H_{23} \\ 0 \end{bmatrix}$$

   that is, the origin $(x, y) = (0, 0)$ 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 $\mathbf{H}$ (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.

   (b) The transformation from $(s, t, 1)$ to points on the rotated 3D plane, and then on to points on the image plane is:

$$\begin{bmatrix} wx \\ wy \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s \\ t \\ 1 \end{bmatrix}$$

   so the homography is:

$$\mathbf{H} = \begin{bmatrix} f & 0 & 0 \\ 0 & f\cos\theta & 0 \\ 0 & \sin\theta & Z_0 \end{bmatrix}.$$

   (c) The points at infinity on the plane are points $\mathbf{H}(s, t, 0)^T$. In the image, these are points (x, y) on the *horizon*. By definition, these are image points $(x, y, 1)$ which (inverse) map to points in the texture space at infinity:

$$\begin{bmatrix} s \\ t \\ 0 \end{bmatrix} \equiv \mathbf{H}^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

The inverse is easily calculated (by hand) to be:

$$\mathbf{H}^{-1} = \begin{bmatrix} \frac{1}{f} & 0 & 0 \\ 0 & \frac{1}{f\cos\theta} & 0 \\ \\ 0 & -\frac{\tan\theta}{Z_0 f} & \frac{1}{Z_0} \end{bmatrix}$$

Which points (x, y) lie on the horizon? From the above reasoning, we must have

$$(0, -\frac{\tan\theta}{Z_0 f}, \frac{1}{Z_0}) \cdot (x, y, 1) = 0$$

and so

$$y = \frac{f}{\tan\theta}.$$

Notice that if the surface is not slanted at all ($\theta = 0$), then $\tan\theta = 0$ and so the horizon is itself the set of points at infinity in the image plane i.e. $(x, y, 0)$.

(d)

$$\begin{aligned} \begin{bmatrix} wx \\ wy \\ w \end{bmatrix} &= \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & X_0 \\ 0 & 0 & h \\ 0 & 1 & Z_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s \\ t \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} f(s\cos\theta - t\sin\theta) + X_0 \\ fh \\ (s\sin\theta + t\cos\theta) + Z_0 \end{bmatrix} \end{aligned}$$

Consider lines that are parallel to either the s or t axes. First, let's get rid of the scale factor w. This gives:

$$(x, y) = (\frac{f(s\cos\theta + t\sin\theta + X_0)}{s\sin\theta - t\cos\theta + Z_0}, \frac{fh}{s\sin\theta - t\cos\theta + Z_0})$$

The two vanishing points defined as follows: Fix s and let $t \to \infty$:

$$\lim_{t\to\infty}(x, y) = (-f\frac{\sin\theta}{\cos\theta}, 0)$$

or, fix t and let $s \to \infty$:

$$\lim_{s\to\infty}(x, y) = (f\frac{\cos\theta}{\sin\theta}, 0).$$

Notice that the vanishing points lie *in the image* along the horizontal line $y = 0$.

Note that varying θ changes the positions of the vanishing points on the horizon. However, changing the origin position (X_0, h, Z_0) within the ground plane (keeping the y coordinate at h) does not change the position of the vanishing points.

8. **lecture 22: essential and fundamental matrix**

   (a) Since $\mathbf{R}_1 = \mathbf{R}_2$, we would simply have $\mathbf{E} = [\mathbf{T}]_\times$, and so

   $$\mathbf{E} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T_X \\ 0 & T_X & 0 \end{bmatrix}$$

   (b) The square of the singular values of an essential matrix $\mathbf{E}$ are the eigenvalues of $\mathbf{E}^T\mathbf{E}$. (This is true for any matrix $\mathbf{A}$, not just an essential matrix – see lecture 18). But $\mathbf{E} = \mathbf{R}_2\mathbf{R}_1^T[\mathbf{T}]_\times$, so

   $$\mathbf{E}^T\mathbf{E} = -[\mathbf{T}]_\times \mathbf{R}_1\mathbf{R}_2^T\mathbf{R}_2\mathbf{R}_1^T[\mathbf{T}]_\times = -[\mathbf{T}]_\times[\mathbf{T}]_\times.$$

   so we just need to know what are the eigenvalues and eigenvectors of $[\mathbf{T}]_\times[\mathbf{T}]_\times$. The transformation $[\mathbf{T}]_\times$ zeros out the component of a vector in the direction of $\mathbf{T}$, that is, $\mathbf{T}$ is a null vector of $[\mathbf{T}]_\times$. So that tells us one eigenvalue (0) and eigenvector ($\mathbf{T}$). What about the other two?

   Multiplying out, we notice that

   $$[\mathbf{T}]_\times^T[\mathbf{T}]_\times = |\mathbf{T}|^2(\mathbf{I} - unit(\mathbf{T})unit(\mathbf{T})^T).$$

   Notice that $\mathbf{I} - unit(\mathbf{T})unit(\mathbf{T})^T$ just zeros the component in direction $\mathbf{T}$, i.e. the eigenvector with eigenvalue 0. For any vector perpendicular to $\mathbf{T}$, multiplying by $[\mathbf{T}]_\times^T[\mathbf{T}]_\times$ just multiples the length by $|\mathbf{T}|^2$. Thus, the two other singular values are $|\mathbf{T}|$.

   (c) The epipole in the first image is $\mathbf{e}_1$ such that $\mathbf{F}\mathbf{e}_1 = \mathbf{0}$, i.e. the null vector of $\mathbf{F}$. If we write $\mathbf{F} = \mathbf{U}\Sigma\mathbf{V}^T$, it follows that $\mathbf{e}_1$ is the third row of $\mathbf{V}^T$ (i.e. the third column of $\mathbf{V}$). The epipole in the second image is $\mathbf{e}_2$ such that $\mathbf{e}_2^T\mathbf{F} = \mathbf{0}$, so $\mathbf{e}_2$ is the third column of $\mathbf{U}$.

9. **lecture 23: stereo correspondence**

   Regions **bc, ij, ef** are binocular. Regions **ab, hi** are seen by left eye only. Regions **cd, fg** are seen by right eye only. Regions **gh, ad** are seen by neither eye.