Geometry of perspective imaging

- Coordinate transformations
- Image formation
- Vanishing points
- Stereo imaging

Images of the 3-D world

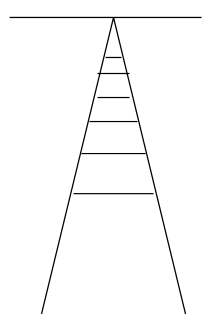
- What is the geometry of the image of a three dimensional object?
  - Given a point in space, where will we see it in an image?
  - Given a line segment in space, what does its image look like?
  - Why do the images of lines that are parallel in space appear to converge to a single point in an image?
- How can we recover information about the 3-D world from a 2-D image?
  - Given a point in an image, what can we say about the location of the 3-D point in space?
  - Are there advantages to having more than one image in recovering 3-D information?
  - If we know the geometry of a 3-D object, can we locate it in space (say for a robot to pick it up) from a 2-D image?

Euclidean versus projective geometry

- Euclidean geometry describes shapes “as they are”
  - properties of objects that are unchanged by rigid motions
    - » lengths
    - » angles
    - » parallelism
- Projective geometry describes objects “as they appear”
  - lengths, angles, parallelism become “distorted” when we look at objects
  - mathematical model for how images of the 3D world are formed

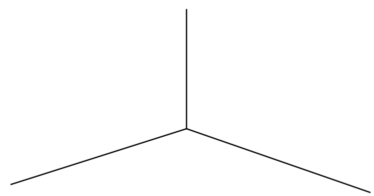
Example 1

- Consider a set of railroad tracks
  - Their actual shape:
    - » tracks are parallel
    - » ties are perpendicular to the tracks
    - » ties are evenly spaced along the tracks
  - Their appearance
    - » tracks converge to a point on the horizon
    - » tracks don’t meet ties at right angles
    - » ties become closer and closer towards the horizon



Example 2

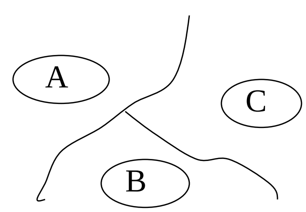
- Corner of a room
  - Actual shape
    - » three walls meeting at right angles. Total of 270° of angle.
  - Appearance
    - » a point on which three lines segments are concurrent. Total angle is 360°

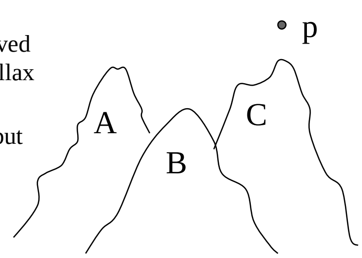


Example 3

q

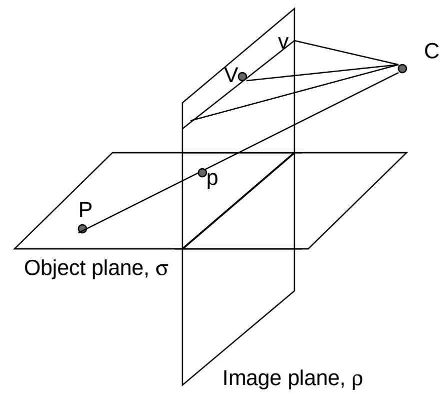
- B appears between A and C from point p
- But from point 1, A appears between B and C
- Apparent displacement of objects due to change in viewing position is called **parallax shift**
  - Greeks knew that is the earth revolved around the sun there would be parallax shift of the stars
  - Tycho Brahe looked for this shift, but his instruments were not accurate enough.





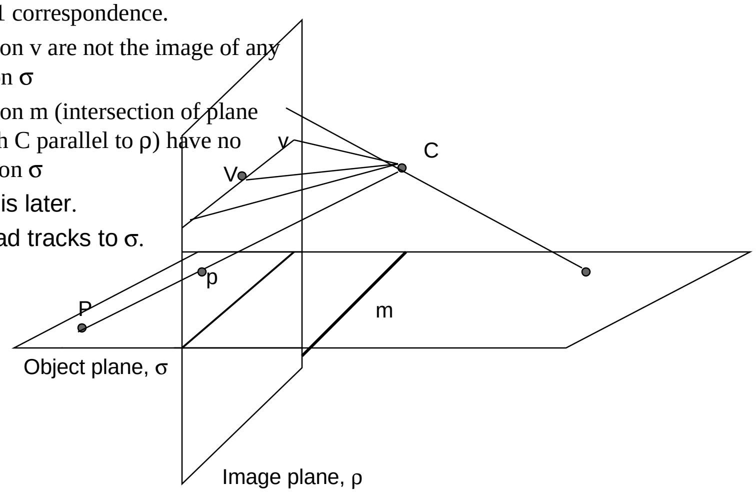
Vanishing points and lines

- To begin with, let's assume we're looking at a scene on the plane, σ, which is perpendicular to the image plane ρ.
  - To each point P on σ we associate the point p on ρ corresponding to the intersection of the line CP with the plane ρ.
  - C is called the center of projection
- The point of intersection of the line through the optical center, C, that is perpendicular to the image plane (and so parallel to the object plane) is called the **principal vanishing point**, V.
- The line, v, which is the intersection of the picture plane and the plane through P parallel to the object plane is called the **vanishing line**, or the **horizon line**.



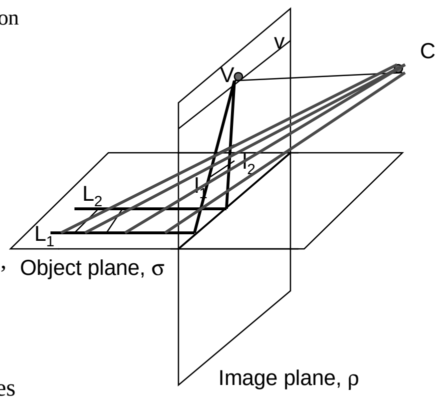
Vanishing points and lines

- Our mapping between σ and ρ does not define a 1-1 correspondence.
  - Points on v are not the image of any point on σ
  - Points on m (intersection of plane through C parallel to ρ) have no image on σ
- We'll fix this later.
- Add railroad tracks to σ.



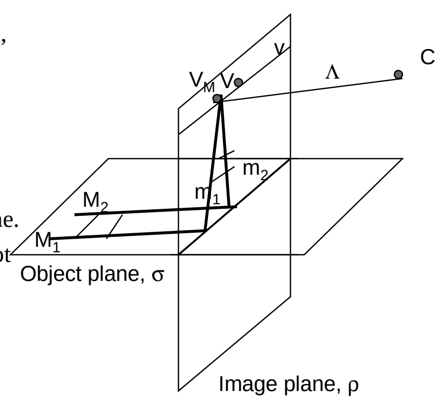
Vanishing points and lines

- The images of all the points on L_1 lie on the plane π_1 defined by L_1 and C.
  - this is because the lines of sight for each point on L_1 lie on this plane.
- The image of L_1 is then the intersection of π_1 with the picture plane - l_1.
- L_1 is parallel to CV since they are both perpendicular to the picture plane.
  - But, then CV must also lie on π_1, since parallel lines are co-planar, and V lies on π_1.
- So, the principal vanishing point must lie on l_1, since it lies on both π_1 and the picture plane.
- Similarly, V must lie on l_2, the image of L_2.
- In fact, ALL lines in the object plane that are perpendicular to the image plane, image to lines passing through V.
  - This is the image of the "point at infinity" for that set of parallel lines



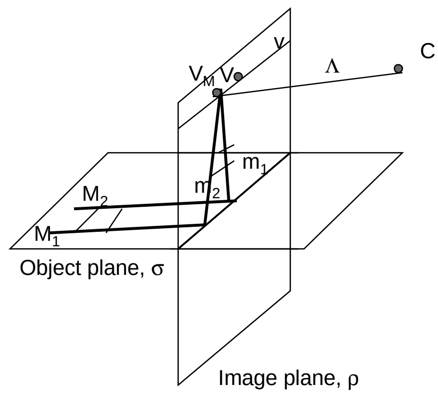
Vanishing points and lines

- M_1 and M_2 are parallel, but not perpendicular to the image plane.
- The image of M_i is the line m_i where the plane, π_i, containing C and M_i intersects the image plane.
- There is a unique line, Λ, through C which is parallel to M_i and so this line is also on π_i.
- Since this line is horizontal, it also lies in the plane determined by C and v, the vanishing line.
- Since v and Λ lie in the same plane, and are not parallel, they must intersect at some point V_m.



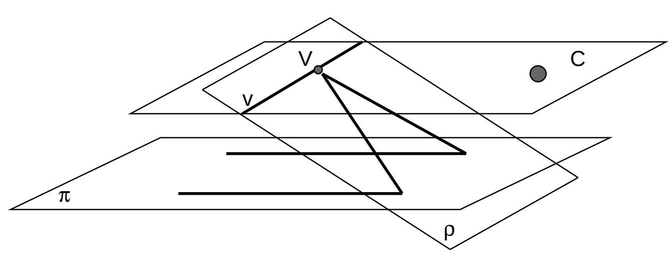
Vanishing points and lines

- This places V_m on both π_i and the image plane, so it must lie on their intersection, m_i.
  - So, any family of parallel lines on the object plane will image to a set of lines passing through a point (vanishing point) on the vanishing line.
- As we rotate the set of lines, the vanishing point moves along the vanishing line.
- So even though parallel lines don't "meet" we can see where they meet in images!

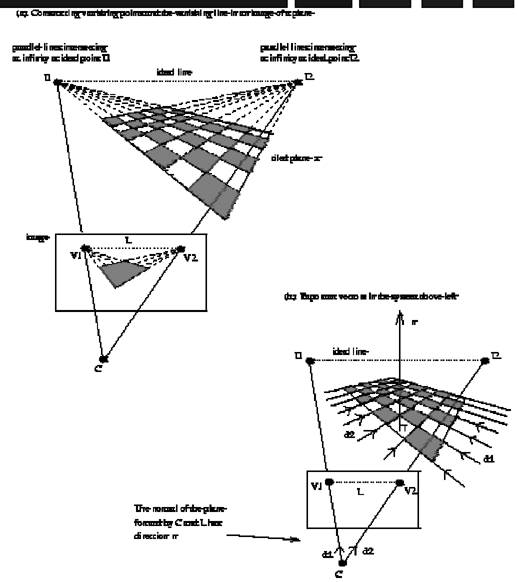


General case

- Generally, the vanishing line for any object plane is the line of intersection of the
  - image plane
  - plane parallel to the object plane through C (called the horizon plane).
- The vanishing point for a line on the object plane is the intersection of the
  - vanishing line
  - plane containing the original line and the center of projection.



Vanishing points and lines



Homogeneous coordinates

- Classical Euclidean geometry: through any point not on a given line, there exists a unique line which is parallel to the given line.
  - For 2,000 years, mathematician tried to “prove” this from Euclid’s postulates.
  - In the early 20’th century, geometry was revolutionized when mathematicians asked: What if this were false?
  - That is, what if we assumed that EVERY pair of lines intersected?
  - To do this, we’ll have to add points and lines to the standard Euclidean plane.
- If (x,y) are the rectangular coordinates of a point, P, and if (x_1, x_2, x_3) are any three real numbers such that:
  - $x_1/x_3 = x$
  - $x_2/x_3 = y$

then (x_1, x_2, x_3) are a set of **homogeneous coordinates** for (x,y).

- So, in particular, (x,y,1) are a set of homogeneous coordinates for (x,y)

Homogeneous coordinates

- Given the homogeneous coordinates, (x_1, x_2, x_3), the rectangular coordinates can be recovered.
- But (x,y) has an infinite number of homogeneous coordinate representations, because if (x_1, x_2, x_3) are homogeneous coordinates of (x,y), then so are $(kx_1, kx_2, k x_3)$ for any k <>0.

Points and lines

- In Euclidean coordinates, we can represent a line by:
  - $l_1 x + l_2 y + l_3 = 0$ $\quad l_1, l_2$ not both 0.
- In homogeneous coordinates, the equation becomes
  - $l_1 x_1 + l_2 x_2 + l_3 x_3 = 0$
- So a line is determined by the three coefficients (l_1, l_2, l_3)
- If l is the line $l_1 x_1 + l_2 x_2 + l_3 x_3 = 0$, then we say that $[l_1, l_2, l_3]$ are the homogeneous coordinates of l.
  - So, in homogeneous coordinates, points and lines look the same - triples of numbers.
  - We'll represent points as (…) and lines as […] to keep things straight.
- A line is uniquely determined by a set of homogeneous coordinates, but a given line has an infinite number of homogeneous coordinate representations.
  - If $[l_1, l_2, l_3]$ represents a given line, then so does $[kl_1, kl_2, k l_3]$.
    - » I.e., the same set of points satisfy the equations $l_1 x_1 + l_2 x_2 + l_3 x_3 = 0$ and $kl_1 x_1 + k l_2 x_2 + kl_3 x_3 = 0$

Points and lines

- We usually think of the equation of a line as specifying the set of points that lie on that line.
  - It's the set of points that satisfies $l_1 x_1 + l_2 x_2 + l_3 x_3 = 0$
- Equivalently, we can think of a point as an equation also! It specifies the set of all **lines** that pass through that point.
  - It's the set of lines that satisfies $x_1 l_1 + x_2 l_2 + x_3 l_3 = 0$
- Now, given two points A = (a_1, a_2, a_3) and B = (b_1, b_2, b_3), what are the coordinates of the line, *l*, that passes through A and B.
  - $a_1 l_1 + a_2 l_2 + a_3 l_3 = 0$
  - $b_1 l_1 + b_2 l_2 + b_3 l_3 = 0$
- The solution is:
  - $l_1 = k(a_2 b_3 - a_3 b_2)$
  - $l_2 = -k(a_1 b_3 - a_3 b_1)$
  - $l_3 = k(a_1 b_2 - a_2 b_1)$
  - for arbitrary k

Points and lines

- The solutions for l_1, l_2, l_3 are the expansions of the second order determinants obtained from the array of the coordinates of A and B

$$\begin{matrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{matrix}$$

  by omitting the first, second and third columns in turn.
  - So, for example:

$$l_1 = k \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} = k(a_2 b_3 - a_3 b_2)$$

Points and lines

- Since points and lines are the "same" it must be true that if the point (x_1, x_2, x_3) is the intersection of $[l_1, l_2, l_3]$ and $[m_1, m_2, m_3]$ then:
  - $x_1 = k(l_2 m_3 - l_3 m_2)$
  - $x_2 = -k(l_1 m_3 - l_3 m_1)$
  - $x_2 = k(l_1 m_1 - l_1 m_1)$
- Now, in rectangular coordinates, if we try to fine the intersection of the following two parallel lines:
  - $l_1 x + l_2 y + l_3 = 0$
  - $l_1 x + l_2 y + l'_3 =0$ $\quad l_3 <> l'_3$
  - we get:

$$x = \frac{\begin{vmatrix} -l_3 & l_2 \\ -l'_3 & l_2 \end{vmatrix}}{\begin{vmatrix} l_1 & l_2 \\ l_1 & l_2 \end{vmatrix}} = \infty(!) \qquad y = \frac{\begin{vmatrix} l_1 & -l_3 \\ l_1 & -l'_3 \end{vmatrix}}{\begin{vmatrix} l_1 & l_2 \\ l_1 & l_2 \end{vmatrix}} = \infty(!)$$

Points and lines

- Solve the same problem in homogeneous coordinates
  - $l_1 x_1 + l_2 x_2 + l_3 x_3 = 0$
  - $l_1 x_1 + l_2 x_2 + l'_3 x_3 = 0$
- Subtracting the equations yields
  - $(l_3 - l'_3)x_3 = 0$
  - since $l_3 <> l'_3$, it must be that $x_3 = 0$
  - so $l_1 x_1 + l_2 x_2 = 0$, which implies that
    - » $x_1 = kl_2$
    - » $x_2 = -kl_1$
- So, the homogeneous point $(kl_2, -kl_1, 0)$ is the intersection of these two lines
  - But, this corresponds to NO point in rectangular coordinates, since its third coordinate is 0.
- So, we just "add" these points to the Euclidean plane, producing what is called the projective plane.
  - For every value of l_1 and l_2, we get one of these points

Points and lines

- Notice that (kl_1, kl_2, 0) lies on <u>every line</u> of the form:
  - $l_1x_1 + l_2x_2 + l_3x_3 = 0$
- So, every family of parallel lines in the Euclidean plane intersects at a single point in the projective plane.
- What are the set of all lines that pass through any of these new points.
  - All of these points have $x_3 = 0$
  - The equation of (a,b,0) in line coordinates is $al_1 + bl_2 + 0l_3 = 0$
  - This line must then be [0,0,1] in homogeneous coordinates
    - » this doesn't correspond to any line in the Euclidean plane, since its first two coordinates are 0.
    - » Every one of our new points lies on this line
    - » So this line is also added to the Euclidean plane
- The points are called **points at infinity**, and the new line is called the **line at infinity**

Perspective imaging - pinhole camera model

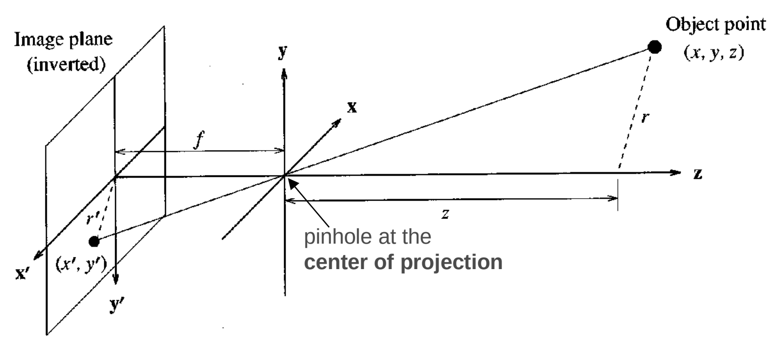
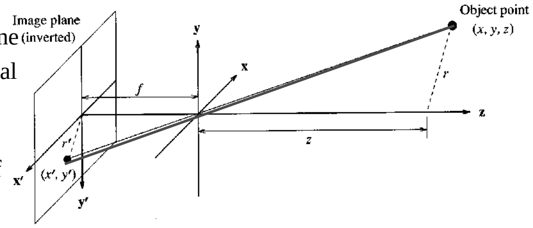


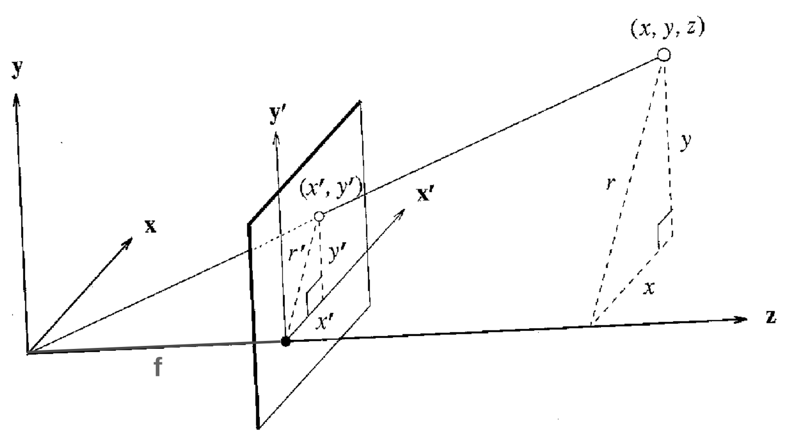
Figure 1.4: The point on the image plane that corresponds to a particular point in the scene is found by following the line that passes through the scene point and the center of projection.

Perspective imaging

- Imaging system center of projection (pinhole)coincides with the origin of the three dimensional coordinate system
- **Line of sight** to a point in the scene is the line through the center of projection to that point
- Image plane is parallel to the x-y plane
  - distance to image plane is f - focal length
  - this inverts the image
  - move the image plane in front of the center of projection

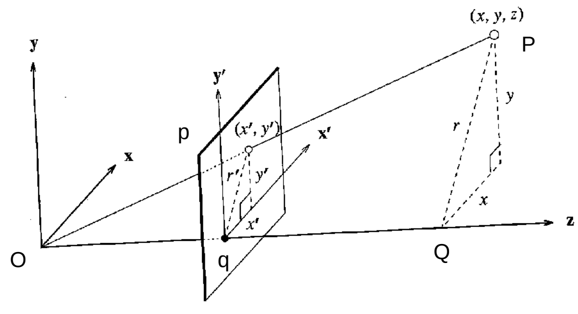


Perspective imaging



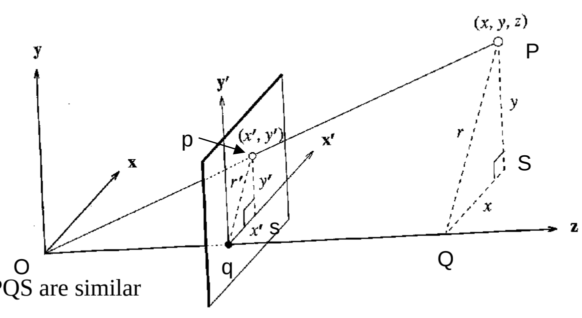
Perspective imaging

- If (x,y,z) is a point in space in front of the camera, what are the coordinates (x',y') of its projection on the image
  - this is the intersection of the line of sight and the image plane
- The distance of (x,y,z) from the z axis is $r = (x^2 + y^2)^{1/2}$
- Consider the following two similar triangles
  - OPQ
  - Opq



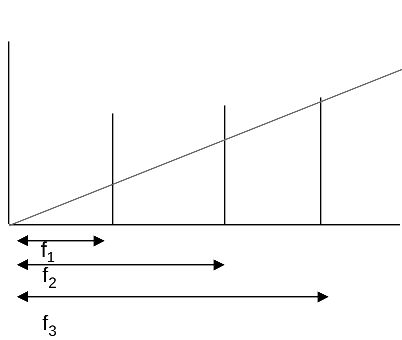
Perspective imaging

- In triangle Opq
  - Oq = f
  - pq = r'
- In triangle OPQ
  - OQ = z
  - PQ = r
- So, f/z = r'/r
- Also, triangles pqs and PQS are similar
  - so x'/x = y'/y = r'/r
- Combining these equations gives
  - x'/x = f/z or x' = fx/z
  - y'/y = f/z or y' = fy/z
- If we set f = 1, then the image of a point (x,y,z) is (x/z, y/z, 1).
  - Note that these are the SAME points in homogeneous coordinates

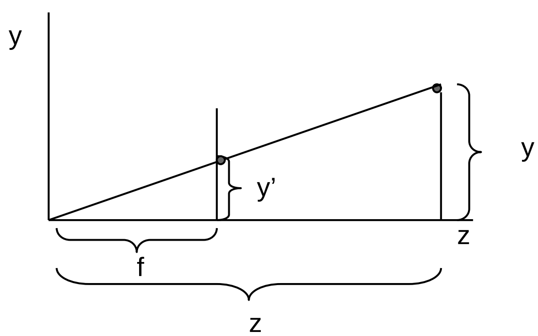


Perspective imaging

- As f gets smaller, image becomes more wide angle (more world points project onto the finite image plane)
- As f gets larger, image becomes more telescopic (smaller part of the world projects onto the finite image plane)



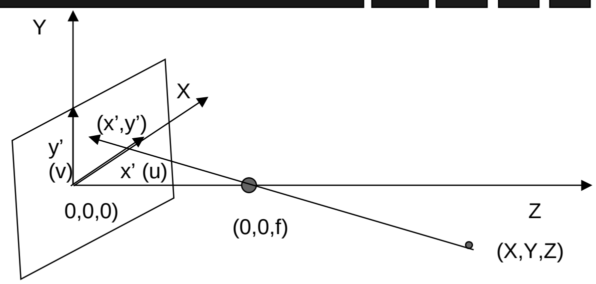
A 2D Example



$$f\,/\,z = y'/\,y$$

$$y' = fy\,/\,z$$

Perspective model



- The plane Z=0 is the **image plane**
  - Sometime referred to as the **(u,v) plane**
- The point(0,0,0) is the **image center**
- The point (0,0,f) is the **focal point**
- The plane Z=f is the **focal plane**
- The point (x',y') is called the **image** or the **perspective projection** of the point (X,Y,Z)

General treatment of vanishing points - Lines in the plane

- In 2-D, we can represent a line by the set of points
  - $L = \{(x,y) | x= a_1 + \lambda b_1, y = a_2 + \lambda b_2\}$
- Examples:
  - $(a_1, a_2) = (0,0)$ and $(b_1, b_2) = (1,1)$
    $x = \lambda$, $y = \lambda$ or $x = y$!
  - $(a_1, a_2) = (0,0)$ and $(b_1, b_2) = (2,2)$
    $x = 2\lambda$, $y = 2\lambda$, so $x=y$ again
    - so, only b_2/b_1 counts - but this is the slope of the line in the plane
  - $(a_1, a_2) = (0,2)$ and $(b_1, b_2) = (1,1)$
    - 45^o line through (0,2)

Properties of perspective imaging

- **3-D straight lines project to 2-D straight lines**
  - let L be a 3-D line

    $L = \{(X,Y,Z)^T \mid \text{for some } \lambda, (X,Y,Z)^T = (a_1,a_2,a_3)^T + \lambda(b_1,b_2,b_3)^T\}$
    - » passes through (a_1,a_2,a_3)
    - » (b_1, b_2, b_3) defines 3-D slope or direction cosines of the line
  - Line L and the point (0,0,f) form a plane, π.
  - The image, ℓ, of L must lie in the intersection of π with the image plane. WHY?
    - » the line from any point, P, on L through (0,0,f) must lie on π because both P and (0,0,f) lie on π. .
    - » The perspective projection of P lies both on this line and on the image plane.
  - Therefore, ℓ must be a line because the intersection of two planes is a line

3D lines to 2D lines

- What is the equation of ℓ?
  - Let π be $Ax + By + Cz = D$
    - » Then the line *l* must have the equation $Ax + By = D$, since it is on the plane Z-0 and the plane π.
  - Since (0,0,f) is on π, $Cf = D$
  - Since (a_1, a_2, a_3) is on π, $Aa_1 + Ba_2 + Ca_3 = Cf$
  - Since π contains ℓ, $Ab_1 + Bb_2 + Cb_3 = 0$
    - » We know that $0 = A(a_1+\lambda b_1) + B(a_2+\lambda b_2) + C(a_3+\lambda b_3) - Cf$ for all λ.
    - » This can be written as
      - $Aa_1 + Ba_2 + Ca_3 - Cf + \lambda(Ab_1 + Bb_2 + Cb_3) = 0$
      - But the red term is 0 from above, so we have
      - $0 + \lambda(Ab_1 + Bb_2 + Cb_3) = 0$ for all λ.
      - Since since this must be true for all λ, the second term must be 0

Properties of perspective imaging

- We can solve for A and B in terms of C

$$A = -C \frac{\begin{vmatrix} a_3 - f & a_2 \\ b_3 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}} \qquad B = -C \frac{\begin{vmatrix} a_1 & a_3 - f \\ b_1 & b_3 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}}$$

Solving for A

- $Aa_1 + Ba_2 + Ca_3 = Cf$
  $Ab_1 + Bb_2 + Cb_3 = 0$
- $Aa_1b_2 + Ba_2b_2 + Ca_3b_2 = Cfb_2$
  $Ab_1a_2 + Bb_2a_2 + Cb_3a_2 = 0$
- $A(a_1b_2 - b_1a_2) + C(a_3b_2 - b_3a_2) = Cfb_2$
- $A = C(b_2(f-a_3) + b_3a_2)/ (a_1b_2 - b_1a_2)$
- $A = -C(b_2(a_3-f) - b_3a_2)/ (a_1b_2 - b_1a_2)$

$$A = -C \frac{\begin{vmatrix} a_3 - f & a_2 \\ b_3 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}}$$

Properties of perspective imaging

- Then ℓ is the line Ax + By = Cf (since z = 0 for all points on L), which after substitution gives:

$$-C\frac{\begin{vmatrix} a_3-f & a_2 \\ b_3 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}}x - C\frac{\begin{vmatrix} a_1 & a_3-f \\ b_1 & b_3 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}}y - Cf = 0$$

$$\begin{vmatrix} a_3-f & a_2 \\ b_3 & b_2 \end{vmatrix}x + \begin{vmatrix} a_1 & a_3-f \\ b_1 & b_3 \end{vmatrix}y + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}f = 0$$

Vanishing points

- Let *L* be a set of straight lines that passes through a point P
  - the images of all L in *L* must pass through the image of P, p.
- (special case) Let P be in the focal plane z=f
  - P = $(a_1, a_2, a_3) = (a_1, a_2, f)$
  - Then the image of any line in *L* is

$$-a_2 b_3 x + a_1 b_3 y + f\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = 0$$

  - They all have slopes a_2/a_1 independent of b_1,b_2,b_3
  - Thus, the images of all the lines in *L* are **parallel** (provided that a_1 and a_2 are not 0 - in which case the lines are in the focal plane and do not have images).
- So, concurrent lines map into concurrent (general case when P is not on the focal plane) or parallel straight lines (special case when P is on the focal plane).

Vanishing points

- Let *L* be a family of **parallel** straight lines of the form:

$L = \{(X,Y,Z)^T \mid$ for some λ, $(X,Y,Z)^T = (a_1,a_2,a_3)^T + \lambda\,(b_1,b_2,b_3)^T \}$

which pass through $(a_1,a_2,a_3)^T$(different for different lines) and have direction cosines $(b_1,b_2,b_3)^T$ (same for all lines)

- The perspective projection $(x',y')^T$ of any point on *L* is given by
  - $x' = f[a_1 + \lambda b_1]/[a_3 + \lambda b_3]$
  - $y' = f[a_2 + \lambda b_2]/[a_3 + \lambda b_3]$
  - For lines which have some slope along the optical axis we must have that b_3 <> 0.
- Points which are on the lines and are infinitely far away from the center of the lens will have the perspective projection

$$x'_\infty = \lim_{\lambda\to\infty} \frac{f\,(a_1 + \lambda b_1)}{a_3 + \lambda b_2} = f b_1 / b_3$$

$$y'_\infty = \lim_{\lambda\to\infty} \frac{f\,(a_2 + \lambda b_2)}{a_3 + \lambda b_2} = f b_2 / b_3$$

Vanishing points

- Since the point (x'_∞, y'_∞) is independent of (a_1,a_2,a_3) all parallel lines which have nonzero slopes along the optic axis have perspective projections which meet at the same **vanishing point**

Vanishing points

- If *L* are parallel to the (x,y) plane, then the argument fails since b_1/b_3 would be undefined ($b_3 = 0$).
  - In this case, L maps into parallel lines in the image.
  - Lines in *L* are of the form $(x-a_1)/b_1 = (y-a_2)/b_2 = 0$ and the image lines are all parallel to $x/b_1 = y/b_2 = 0$ in the image plane
- Example: Family of lines parallel to the y-axis: $b_1=b_3=0$, $b_2=1$. Vertical lines all map into vertical lines (x = constant) in the image plane

Vanishing points

- Example: Family of lines perpendicular to the y-axis - i.e., horizontal lines parallel to the x-z plane so that $b_2= 0$ (can rotate them around the z axis)
  - vanishing point is on the x axis: $x = -b_1f/b_3$, $y = 0$
  - as the lines become parallel to the z axis, the vanishing point approaches 0 (b_1 approaches 0)
  - as they become parallel to the x axis (b_3 approaches 0) it recedes to infinity
  - generally, as we rotate the set of parallel lines on a plane, their vanishing point moves along a line in the image called the horizon line (or vanishing line) for that plane.

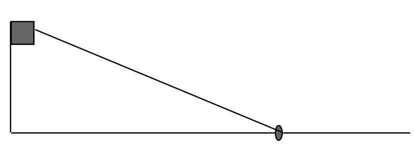
Importance of vanishing points

- From the position of the vanishing point in the image we can recover the direction cosines of the 3D parallel lines that meet at that vanishing point
-

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \frac{1}{\sqrt{x'^2_\infty + y'^2_\infty + f^2}} \begin{pmatrix} x'_\infty \\ y'_\infty \\ f \end{pmatrix}$$

Point in a known plane

- Motivation: robot driver
  - Suppose a robot is driving on a flat road
  - It “sees” a turn in the road in the image
  - How far down the road, in 3-D is this turn.
- Assume robot knows the equation of the road plane in the (X,Y,Z) coordinate system attached to the camera.
  - then the 3-D position of any point on the road can be determined.



Point in a known plane

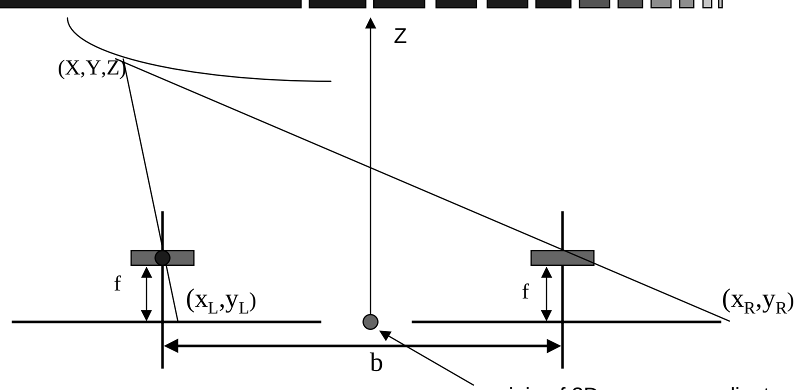
- Let $(a\ b\ c)^T$ be the 3-D point and $(u\ v)^T$ be its perspective projection. Then
  - u= fa/c
  - v= fb/c
- Let the ground plane be $\{(X,Y,Z)^T \mid AX + BY + CZ + D = 0\}$
- Since the point lies on the plane, Aa + Bb + Cc + D = 0
- So, A(uc/f) + B(vc/f) + Cc + D = 0, using the projection equations above. This equation can be solved for c, and then the projection equations can be used to solve for a and b

$$c = \frac{-Df}{Au + Bv + Cf}$$

$$b = \frac{-Dv}{Au + Bv + Cf}$$

$$c = \frac{-Du}{Au + Bv + Cf}$$

Stereo imaging

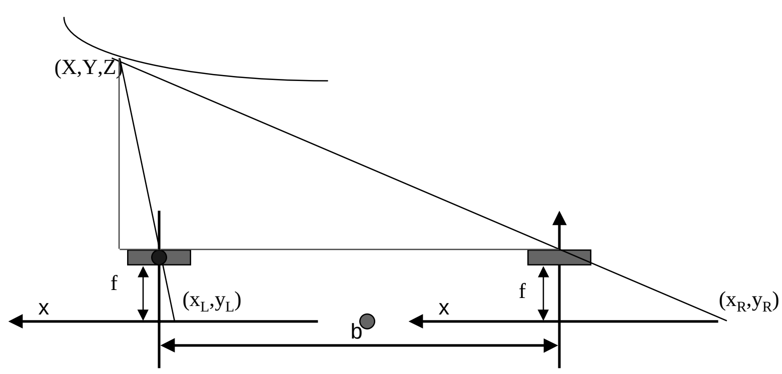


- Optical axes are parallel
- optical axes separated by **baseline**, b.
- line connecting lens centers is perpendicular to the optical axis, and the x axis is parallel to that line
- 3D coordinate system is a **cyclopean** system centered between the cameras

Stereo imaging

- (X,Y,Z) are the coordinates of P in the Cyclopean coordinate system.
- The coordinates of P in the left camera coordinate system are $(X_L, Y_L, Z_L) = (X-b/2, Y, Z)$
- The coordinates of P in the right camera coordinate system are $(X_R, Y_R, Z_R) = (X+b/2, Y, Z)$
- So, the x image coordinates of the projection of P are
  - $x_L = (X+b/2)f/Z$
  - $x_R = (X-b/2)f/Z$
- Subtracting the second equation from the first, and solving for Z we obtain:
  - $Z = bf/(x_L - x_R)$
- We can also solve for X and Y:
  - $X = b(x_L + x_R)/2(x_L-x_R)$
  - $Y = by/(x_L-x_R)$

Stereo imaging

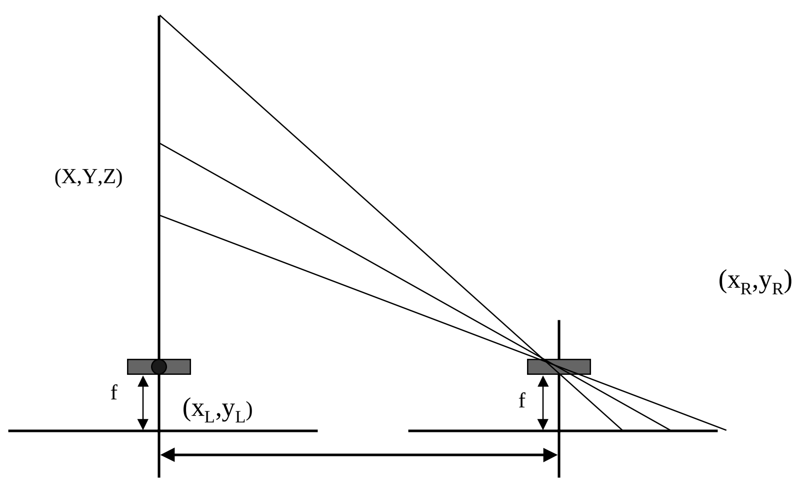


- $x_L - x_R$ is called the **disparity**, d, and is always negative
- $X = (b[x_R + x_L]/2)/ d$ $Y = by/ d$ $Z = bf/d$

Stereo imaging

- Distance is inversely proportional to |disparity|
  - disparity of 0 corresponds to points that are infinitely far away from the cameras
  - in digital systems, disparity can take on only integer values (ignoring the possibility of identifying point locations to better than a pixel resolution)
  - so, a disparity measurement in the image just constrains distance to lie in a given range
- Disparity is directly proportional to b
  - the larger b, the further we can accurately range
  - but as b increases, the images decrease in common field of view

Range versus disparity



Stereo imaging

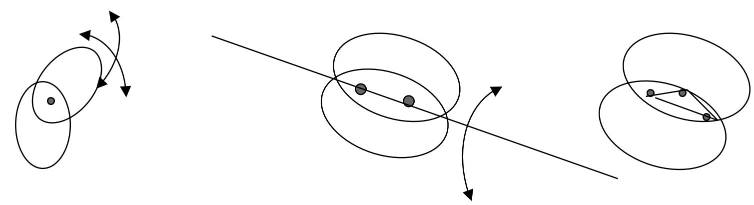
- Definition: A scene point, P, visible in both cameras gives rise to a pair of image points called a **conjugate pair**.
  - the conjugate of a point in the left (right) image must lie on the same image row (line) in the right (left) image because the two have the same y coordinate
  - this line is called the **conjugate line**.
  - so, for our simple image geometry, all conjugate lines are parallel to the x axis

A more practical stereo image model

- Difficult, practically, to
  - have the optical axes parallel
  - have the baseline perpendicular to the optical axes
- Also, we might want to tilt the cameras towards one another to have more overlap in the images
- Calibration problem - finding the transformation between the two cameras
  - it is a rigid body motion and can be decomposed into a rotation, **R**, and a translation, **T**.

A more practical stereo image model

- Let $p_L = (X_L, Y_L, Z_L)^T$ be the position of P in the left system
- Let $p_R = (X_R, Y_R, Z_R)^T$ be the position of P in the right system
- Then $p_R = Rp_L + T$ where R is a 3x3 orthonormal matrix representing the rotation ($RR^T = I$) and T is the translation vector
- This represents a set of 3 linear equations in 12 unknowns; so a set of 4 non-coplanar pairs of 3-D points is sufficient, in general to solve for the transformation
  - in fact, orthonormality of R means only 3 points are required



A more practical stereo image model

- Generally, we only have the conjugate pairs, not the 3-D coordinates of the scene points in the two coordinate systems
- For a given conjugate pair, we have:

$$r_{11}x_L + r_{12}y_L + r_{13} f + t_1f/Z_L = x_RZ_R/Z_L$$
$$r_{21}x_L + r_{22}y_L + r_{23} f + t_2f/Z_L = y_RZ_R/Z_L$$
$$r_{31}x_L + r_{32}y_L + r_{33} f + t_3f/Z_L = fZ_R/Z_L$$

- This is 3 equations in 14 unknowns
- Each additional conjugate pair adds 3 equations, but also adds two unknowns (the Z coordinates)
- For n points we have 3n equations and 12 + 2n unknowns. So, we need at least 12 points to satisfy the equations

A more practical stereo image model

- Note that the unknowns t_i and Z_{Lj}, Z_{Rj} appear only in ratios
  - So, if **T**, Z_{Lj}, Z_{Rj} is a solution, then so is k**T**, kZ_{Lj}, kZ_{Rj} for any k not equal to 0
  - This means that absolute range cannot be determined - must have one more constraint to get a unique solution (e.g., (**T•T**)=1

General stereo matching

- Assume relative orientation of cameras is known
- An image point (x_L, y_L) in the left coordinate system is the image of some point on a ray through the origin of the left camera coordinate system. The points on this ray all have coordinates of the form

  $X_L = x_L s$ $Y'_L = y_L s$ $Z_L = fs$ since all points of this form project onto (x_L, y_L)
- In the right image system, the coordinates of points on this ray are:

  $X_R = (r_{11}x_L + r_{12}y_L + r_{13} f)s + t_1$

  $Y_R = (r_{21}x_L + r_{22}y_L + r_{23} f)s + t_2$

  $Z_R = (r_{31}x_L + r_{32}y_L + r_{33} f)s + t_3$
- These points project onto $x_R/f = X_R/Z_R$ and $y_R/f = Y_R/Z_R$

General stereo matching

- Let

  $X_R = as + u$

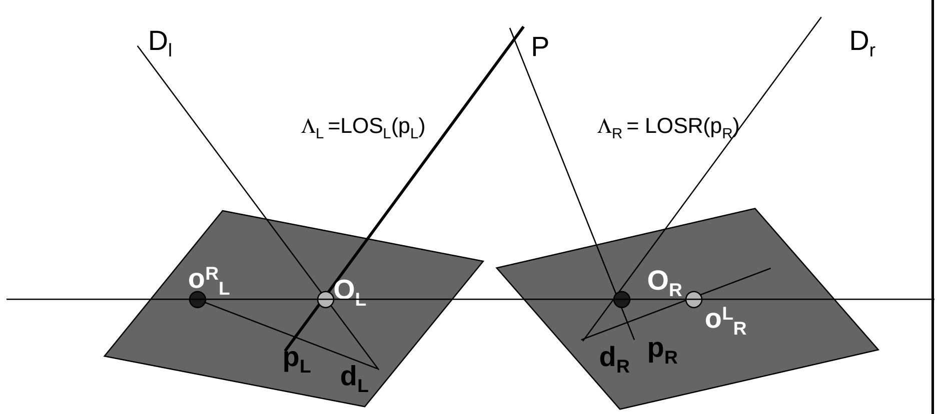
  $Y_R = bs + v$

  $Z_R = cs + w$
- Then

  $x_R/f = (as+u)/(cs+w)$

  $y_R/f = (bs+v)/(cs+w)$
- This is a straight line connecting the point
  - (u/w, v/w) which occurs for s=0 and is the image of the left camera center in the right camera coordinate system, to
  - (a/c, b/c) which occurs as s approaches infinity,the vanishing point for the ray

General stereo geometry



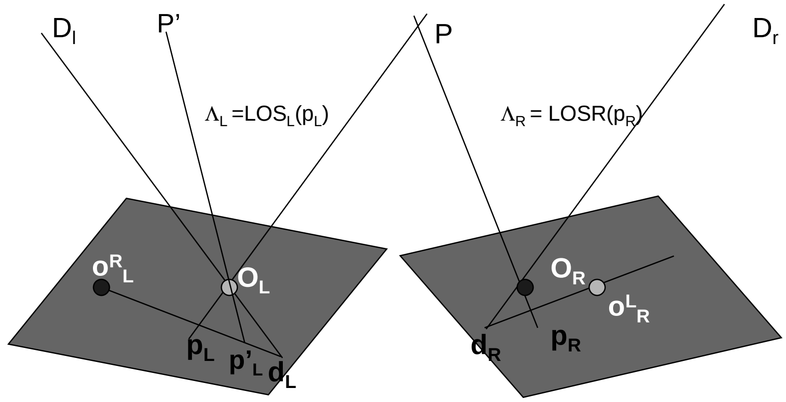
General stereo imaging

- **FACTS**
  - Point P lies somewhere on the ray (line) Λ_L from p_L through O_L
    - but from the left image alone, we do not know where on this ray P lies
  - Since the perspective projection of a line is a line, the perspective projection of Λ_L in the right image is a line
    - the “first” point on Λ_L that might correspond to P is O_L
      - any point closer to the left image than O_L would be between the lens and the image plane, and could not be seen
      - the perspective projection of O_L in the right camera is the point o^L_R.
    - the “last” point on Λ_L that might correspond to P is the point “infinitely” far away along the ray Λ_L
      - but its image is the vanishing point of the ray Λ_L in the right camera, d_R
    - any other possible location for P will project to a point in R on the line joining o^L_R to d_R.

General stereo imaging

- **First general conclusion**
  - Given any point, p_L, in the left image of a stereo pair, its **conjugate point** must appear on a **line** in the right image
  - Furthermore, all of the conjugate lines for all of the points in the left image must pass through a common point in the right image
    - » this is image of the left lens center in the right image
    - » this point lies on the line of sight for **every** point in the left image
    - » therefore, the conjugate lines must all contain (i.e., pass through) the image of this point
    - » This point is called an **epipole.**
  - Finally, the conjugate line for p_L must also pass through the vanishing point in the right image for the line of sight through p_L
  - Since we know, based on p_L’s coordinates, the coordinates in the right image of two points on its conjugate line, we know its conjugate line - **BUT THERE IS ONE MORE POINT!**

General stereo imaging



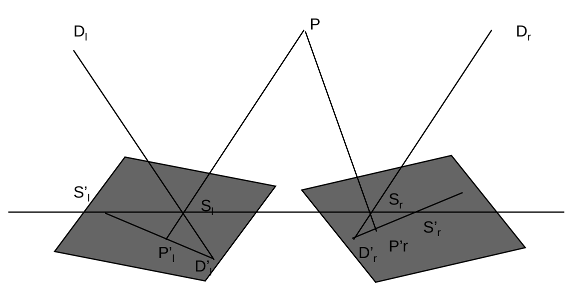
Next **FACT**

- Remember that
  - any three non-collinear points define a plane, and
  - the intersection of two planes is a straight line
- The points O_L, p_L, and o^R_L are three non-collinear points, so they form a plane, Π
  - the line Λ_L lies on this plane, since two points on the line lie on the plane
- The intersection of this plane with the right image plane is the conjugate line of p_L
  - and this would be the image of **any** line on this plane
- Let p'_L be some other point on the line joining p_L and o^R_L.
  - the line of sight through p'_L to P' lies on Π since two points on that line (p_L and o^R_L) lie on the plane
- ***Therefore***, the conjugate line for p'_L must be the same line as the conjugate line for $p_{L,}$ or for **any other point** on the line containing p_L and o^R_L.
- The lines p_L - o^R_L and p_R - o^R_L are called **epipolar lines**

Second general conclusion

- Given any point, p_L, in the right image
  - it lies on a line containing the image of the right camera center in the left image, and
  - it has a conjugate line in the right camera
- Given **any point** on either of these two lines, its conjugate pair must lie on the other line
- These lines are called epipolar lines.

General stereo geometry



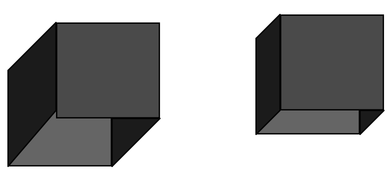
- Points S'_l, P'_l, and D'_l form a straight line
  - They all lie on the image of the line containing S_r, P and that line's vanishing point
  - There is a similar line in the right image
- These lines are called **conjugate** or **epipolar** lines and constrain matching
- All epipolar lines pass through one point - the image of the other lens center (which might not be in the finite field of view)

Stereo correspondence problem

- Given a point,p, in the left image, find its conjugate point in the right image
  - called the stereo correspondence problem
  - solve using either gray level correlation or edge correlation
- What constraints simplify this problem?
  - Epipolar constraint - need only search for the conjugate point on the epipolar line
  - Positive disparity constraint - need only search the epipolar line to the “right” of the vanishing point in the right image of the ray through p in the left coordinate system
  - Continuity constraint - if we are looking at a continuous surface, images of points along a given epipolar line will be ordered the same way

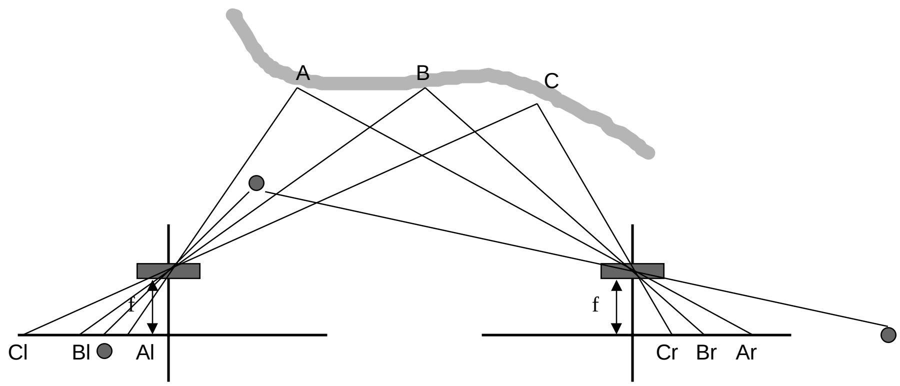
Stereo correspondence problem

- Similarity of correspondence functions along adjacent epipolar lines
- Disparity gradient constraint - disparity changes slowly over most of the image.
  - Exceptions occur at and near occluding boundaries where we have either discontinuities in disparity or large disparity gradients as the surface recedes away from sight.



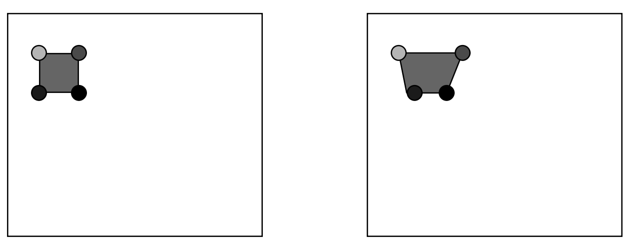
As the blue surfaces become more slanted, they occupy a smaller area of the image

Continuity constraint



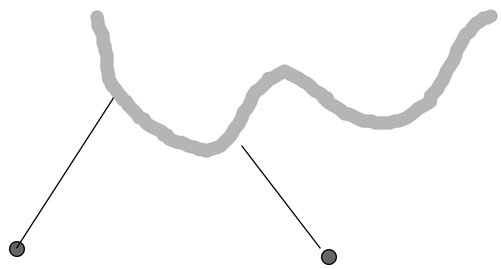
Why is the correspondence problem hard?

- Foreshortening effects
  - Match images at low resolutions
  - User resulting disparity map to warp images
  - Match at next higher resolution
- A square match window in one image will be distorted in the other if disparity is not constant - complicates correlation

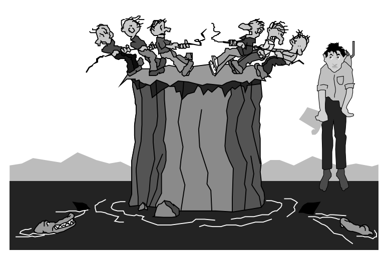


Why is the correspondence problem hard

- Occlusion
  - Even for a smooth surface, there might be points visible in one image and not the other
  - Consider aerial photo pair of urban area - vertical walls of buildings might be visible in one image and not the other
  - scene with depth discontinuities (lurking objects) violate continuity constraint and introduces occlusion



Why is the correspondence problem hard?

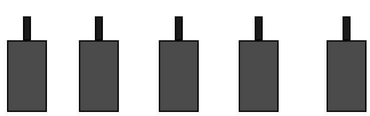


Why is the correspondence problem hard?

- Variations in intensity between images due to
  - noise
  - specularities
  - shape-from-shading differences
- Coincidence of edge and epipolar line orientation
  - consider problem of matching horizontal edges in an ideal left right stereo pair
  - will obtain good match all along the edge
  - so, edge based stereo algorithms only match edges that cross the epipolar lines

Multi-camera stereo

- Slider stereo
  - single baseline
  - equal separation of camera
  - can track features from left to right or vote on best match



Multi camera stereo

- multiple baseline systems
  - camera pairs have different epipolar geometry
  - must find matches in consistent positions in most or all images
  - allows matching of edges in all orientations