Solutions to Chapter 13

Prob. 13.1 — Define retinal disparity. Suppose your two eyes are fixated on a small object at some distance away. What is the retinal disparity of the object?

Answer (Prob. 13.1) — Retinal disparity is the relative shift between the locations on the two retina on which a scene point is projected. If the eyes are fixated on an object, the retinal disparity of the object is zero, since the object projects onto the center of both retinas.

Prob. 13.2 — Suppose two cameras are positioned such that both image planes are parallel, but one plane is slightly in front of the other. Are the cameras rectified? Why or why not?

Answer (Prob. 13.2) — No the cameras are not rectified. In the general case (e.g., if the cameras have the same focal length), then the epipolar lines will not be the scanlines of the images. However, in the special case that the focal lengths are such that the baseline is parallel to the image planes, the epipolar lines may in fact be the scanlines, but a scaling factor will distinguish the corresponding scanlines, so the cameras will still not be rectified.

Prob. 13.3 — Suppose an object 2 m away is viewed by a rectified pair of stereo cameras with a baseline of 50 mm, and the lens of each camera has a focal length of 35 mm. What is the disparity?

Answer (Prob. 13.3) — The disparity is given by $d = fb/z = (35)(50)/2000 = 1750/2000 = 0.875$ mm.

Prob. 13.4 — Which of the following pixels in the right image could possibly match the pixel $(52, 3)$ in the left image, assuming the images are rectified, and the maximum disparity is 20? (a) $(26, 3)$, (b) $(48, 13)$, (c) $(64, 3)$, (d) $(48, 3)$, and (e) $(59, 6)$.

Answer (Prob. 13.4) — Only the pixel in (d) could match the pixel. Note that (b) and (e) are not possible since they are on the wrong row; (c) is not possible due to the cheirality constraint ($i.e., 64 > 52$); and (a) exceeds the maximum disparity of 20 ($i.e., 52 - 26 = 26 > 20$).

Prob. 13.5 — Explain the relationship between the epipolar constraint and rectified cameras.

Answer (Prob. 13.5) — The epipolar constraint applies to any pair of stereo cameras. In the special case that the pair is rectified, the epipolar constraint implies that the epipolar lines are the scanlines of the images.

Prob. 13.6 — Consider the ordering constraint. (a) What other constraint is implied by it? (b) What zone describes the set of matches that it forbids? (c) What is another name for the constraint? (d) Give an example when it is violated.

Answer (Prob. 13.6) — (a) The ordering constraint implies the uniqueness constraint. (b) The ordering constraint forbids matches within the forbidden zone. (c) The ordering constraint is also known as the monotonicity constraint. (d) The ordering constraint is violated when a thin object (i.e., thinner than the baseline) is close to the camera (relative to the background).

Prob. 13.7 — Consider a thin, opaque pole just thicker than the interpupillary distance. How close does the pole have to be to the cameras in order to violate the ordering constraint?

Answer (Prob. 13.7) — It is impossible to violate the ordering constraint with such a pole. See Figure 13.8. The thickness of the pole would have to be thinner than the interpupillary distance for it to be even possible for the ordering constraint to be violated.

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Prob. 13.8 — The function $d^*(x) \equiv \max_a(d(x+a) - \kappa|a|)$ is equivalent to grayscale dilation of the function d with what 1D structuring element?

Answer (Prob. 13.8) — This function is equivalent to grayscale dilation of the function d with a triangle function whose base and height both equal $2a$, which is easily verified graphically.

Prob. 13.9 — Given a constant κ, a function f is said to be *Lipschitz continuous* if and only if

$$|f(x+h) - f(x)| \leq \kappa|h| \tag{13.202}$$

for all x and h. The smallest such κ is called the Lipschitz constant of the function, and the function is called a Lipschitz function. Lipschitz continuity is a smoothness condition on functions that is stronger than regular continuity. Show that, if the disparity gradient limit is satisfied, then the Cyclopean disparity function is Lipschitz continuous with the same constant κ.

Answer (Prob. 13.9) — The Cyclopean disparity function is $\partial d/\partial x$. If the disparity gradient limit is satisfied, then the Cyclopean disparity function satisfies $|\partial d/\partial x| \leq \kappa$ for some constant κ. From the definition of derivative, $|\lim_{h\to 0}(d(x+h) - d(x))/h| \leq \kappa$. Since this is true for all x, it implies $|(d(x+h) - d(x))/h| \leq \kappa$. Rearranging yields $|(d(x+h) - d(x))| \leq \kappa|h|$.

Prob. 13.10 — Suppose the left disparity map of a scanline from a pair of rectified images is given by $0, 0, 0, 2, 2, 2, 0, 0, 0$ for the pixels $x_L = 0$ through $x_L = 8$.

(a) Compute the right disparity map for $x_R = 0$ through $x_R = 8$, assuming that occluded pixels are part of the background.

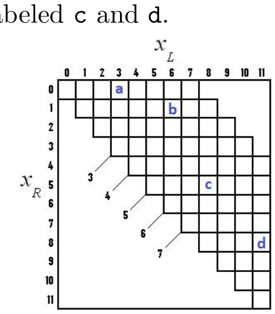
(b) Compute the Cyclopean disparity function. Be sure to label the coordinates x of the points, and do not restrict the computation to integer coordinates.

Answer (Prob. 13.10) — (a) The right disparity map is given by $0, 2, 2, 2, 0, 0, 0, 0, 0$. (b) Each row in the table below indicates a match:

x_L	x_R	$x = \frac{1}{2}(x_L + x_R)$	d(x)
0	0	0	0
3	1	2	2
4	2	3	2
5	3	4	2
6	6	6	0
7	7	7	0
8	8	8	0

The Cyclopean disparity function is given by the last column.

Prob. 13.11 — Compute the disparity gradient (a) between the matches labeled `a` and `b` in the figure below, and (b) between the matches labeled `c` and `d`.



Answer (Prob. 13.11) — The answers are as follows:

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

(a) Disparity gradient between a and b: $= (5-3)/(3.5-2.5) = 2$.

(b) Disparity gradient between c and d: $= (3-3)/(9.5-6.5) = 0$.

Prob. 13.12 — Implement a basic algorithm for matching two rectified stereo images.

(a) Implement BlockMatch1 of Algorithm 13.1.

(b) Implement BlockMatch2 of Algorithm 13.2, and compare running times for different values of w.

(c) Add the left-right check using Algorithm 13.4, and compare results with those of (a) and (b). What do you notice?

Answer (Prob. 13.12) — Answers may vary. The left-right check discards some of the pixels (most notably, those near object boundaries).

Prob. 13.13 — Show that minimizing the SSD is nearly the same as maximizing the cross-correlation. Under what circumstances are they identical?

Answer (Prob. 13.13) — Expanding the SSD term,

$$[I_L(\mathbf{x}_L) - I_R(\mathbf{x}_R)]^2 = [I_L(\mathbf{x}_L)]^2 + [I_R(\mathbf{x}_R)]^2 - 2I_L(\mathbf{x}_L)I_R(\mathbf{x}_R). \tag{13.203}$$

Recall that in the block matching algorithm, the window in the left image is fixed, and a search is conducted for the best disparity in the right image. Therefore,

$$\begin{aligned}\min_d \sum_{\mathbf{x}_L \in \mathcal{W}} [I_L(\mathbf{x}_L) - I_R(\mathbf{x}_L + d)]^2 &= \min_d \sum_{\mathbf{x}_L \in \mathcal{W}} [I_L(\mathbf{x}_L)]^2 + [I_R(\mathbf{x}_L + d)]^2 - 2I_L(\mathbf{x}_L)I_R(\mathbf{x}_L + d) \\ &= \max_d \sum_{\mathbf{x}_L \in \mathcal{W}} -[I_L(\mathbf{x}_L)]^2 - [I_R(\mathbf{x}_L + d)]^2 + 2I_L(\mathbf{x}_L)I_R(\mathbf{x}_L + d).\end{aligned}$$

where $\mathcal{W}$ is the window, and the notation $\mathbf{x}_L + d$ is shorthand for $(x_L + d, y_L)$. Since the first term on the right side is not dependent upon d, it does not affect the solution for the best disparity:

$$\begin{aligned}\arg\min_d \sum_{\mathbf{x}_L \in \mathcal{W}} [I_L(\mathbf{x}_L) - I_R(\mathbf{x}_L + d)]^2 &= \arg\max_d \sum_{\mathbf{x}_L \in \mathcal{W}} -[I_R(\mathbf{x}_L + d)]^2 + 2I_L(\mathbf{x}_L)I_R(\mathbf{x}_L + d) \\ &\approx \arg\max_d \sum_{\mathbf{x}_L \in \mathcal{W}} I_L(\mathbf{x}_L)I_R(\mathbf{x}_L + d),\end{aligned}$$

where the second line arises from the assumption that the energy $[I_R(\mathbf{x}_L + d)]^2$ does not vary much across the image. If the energy in all the overlapping windows in the right image are the same (i.e., statistically speaking, the image is *stationary*), then the approximation becomes an equality, in which case minimizing the SSD is identical to maximizing the cross-correlation.

Prob. 13.14 — Compute the rank and census transforms of the following window of grayscale pixel values:

$$\begin{bmatrix} 8 & 5 & 3 \\ 7 & 6 & 9 \\ 1 & 1 & 2 \end{bmatrix}$$

Answer (Prob. 13.14) — The census transform applies the Heaviside operator to each pixel in the window after subtracting the value of the central pixel. This indicates which pixels have value greater

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

than or equal to that of the central pixel:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The rank transform is simply the number of pixels whose value is less than that of the central pixel (i.e., the number of zeros above), namely 5.

Prob. 13.15 — Apply Algorithm 13.5 to compute the edit distance, as well as the matching function, between the strings "cart" and "earth".

Answer (Prob. 13.15) — By inspection, the matching function matches the 4 letters in "cart" to the 4 letters in "eart" with 'h' having no match; from which we see that the edit distance is 2 (i.e., 1 point for changing 'c' to 'e' and 1 point for adding 'h'). These results are verified by applying the algorithm, which fills in the two tables below with the costs and previous matches, respectively, where $\emptyset$ is the empty string:

Φ table:

	$\emptyset$	e	a	r	t	h
$\emptyset$	0	1	2	3	4	5
c	1	1	2	3	4	5
a	2	2	1	2	3	4
r	3	3	2	1	2	3
t	4	4	3	2	1	2

Π table:

	$\emptyset$	e	a	r	t	h
$\emptyset$	$\cdot$	$\leftarrow$	$\leftarrow$	$\leftarrow$	$\leftarrow$	$\leftarrow$
c	$\uparrow$	$\Nwarrow$	$\leftarrow\nwarrow$	$\leftarrow\nwarrow$	$\leftarrow\nwarrow$	$\leftarrow\nwarrow$
a	$\uparrow$	$\uparrow\nwarrow$	$\Nwarrow$	$\leftarrow$	$\leftarrow$	$\leftarrow$
r	$\uparrow$	$\uparrow\nwarrow$	$\uparrow$	$\Nwarrow$	$\leftarrow$	$\leftarrow$
t	$\uparrow$	$\uparrow\nwarrow$	$\uparrow$	$\uparrow$	$\Nwarrow$	$\Leftarrow$

The double arrows indicate the matching function between the two strings. (Cells with more than one arrow indicate a tie.)

Prob. 13.16 — Explain the key idea behind semi-global matching.

Answer (Prob. 13.16) — Semi-global matching runs dynamic programming along multiple directions within the image (not necessarily along the scanlines) to overcome the streaks that occur when applying dynamic programming in the traditional fashion.

Prob. 13.17 — Explain the difference between motion field and optical flow. Given an example where a nonzero motion field results in zero optical flow, and vice versa.

Answer (Prob. 13.17) — The motion field is the 2D projection of the actual 3D velocity vectors of scene points, whereas optical flow consists of the 2D vectors of *apparent* velocity. A nonzero motion field results in zero optical flow when the surface has no texture, for example, a rotating ping pong ball. Nonzero optical flow results from zero motion field when the motion is only apparent, for example, a movie projected onto a wall at the cinema.

Prob. 13.18 — List some situations that would cause the brightness constancy assumption to fail.

Answer (Prob. 13.18) — The brightness constancy assumption fails when the lighting changes from frame to frame, the object move into or out of shadow, the automatic gain control of the camera changes, pixels near an occluding boundary corrupt the values within the window, and so forth.

Prob. 13.19 — Explain (a) why the optical flow constraint equation in Equation (13.28) is actually an approximation, not an equation. Also explain (b) why it contains a mixture of partial and total derivatives, and (c) how $\partial I/\partial t$ is usually computed.

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Answer (Prob. 13.19) — (a) The optical flow constraint equation is an approximation because it arises from a linear Taylor series, that is, the Taylor series truncated to the linear term (ignoring higher-order terms). (b) Since x and y are functions only of t, their derivatives are total derivatives; whereas the image I is a function of x, y, and t, which explains why its derivatives are partial derivatives. (c) the partial derivative of the image with respect to time is usually computed by subtracting consecutive images pixelwise.

Prob. 13.20 — Implement the Lucas-Kanade method.

Answer (Prob. 13.20) — Answers may vary.

Prob. 13.21 — What problem is generalized Lucas-Kanade attempting to solve?

Answer (Prob. 13.21) — Generalized Lucas-Kanade attempts to match two image patches under more complex motion models than simple translation.

Prob. 13.22 — Explain how the Horn-Schunck algorithm relates to the Lucas-Kanade method. What are some similarities and differences?

Answer (Prob. 13.22) — Horn-Schunck computes the optical flow at *every* pixel in the image by assuming that neighboring pixels yield *similar* optical flow values. Lucas-Kanade, on the other hand, computes the optical flow at *some* pixels in the image by assuming that neighboring pixels yield the *same* optical flow values. Both methods are used to estimate the optical flow.

Prob. 13.23 — Suppose we set $\lambda = 0$ in the Horn-Schunck algorithm. (a) What is the rank of the matrix on the left-hand side of Equation (13.77), and what are the implications for solving for $\begin{bmatrix} u & v \end{bmatrix}^\mathsf{T}$? (b) State what direction $\begin{bmatrix} u & v \end{bmatrix}^\mathsf{T}$ will be shifted relative to $\begin{bmatrix} \bar{u} & \bar{v} \end{bmatrix}^\mathsf{T}$, and relate this answer with the finding of Equation (13.31).

Answer (Prob. 13.23) — (a) The rank is 1, meaning that it is not possible to solve for $\begin{bmatrix} u & v \end{bmatrix}^\mathsf{T}$. This is because with $\lambda = 0$ there is no regularization, and regularization is needed to overcome the fact that the optical flow constraint equation is underdetermined. (b) As seen in Equation (13.83), the flow vector is shifted relative to the average flow vector of the previous iteration by the gradient of the image, that is, along the direction ∇I. This result agrees with Equation (13.31), which states that only the component of motion in the direction of the gradient can be computed.

Prob. 13.24 — Given the point $(5, 2)$ in the 2D Euclidean plane, (a) write the homogeneous coordinates of the point, (b) write the coordinates of the equivalent point on the $w = 1$ plane, and (c) write the coordinates of the equivalent point on the positive unit hemisphere.

Answer (Prob. 13.24) — (a) The homogeneous coordinates are $(5, 2, 1)$. The point on the $w = 1$ plane is $\begin{bmatrix} 5 & 2 & 1 \end{bmatrix}^\mathsf{T}$. (b) The equivalent point on the positive unit hemisphere is $\frac{1}{\sqrt{30}} \begin{bmatrix} 5 & 2 & 1 \end{bmatrix}^\mathsf{T}$, because $\sqrt{5^2 + 2^2 + 1^2} = \sqrt{30}$.

Prob. 13.25 — Convert the homogeneous coordinates $(27, 18, 3)$ to inhomogeneous coordinates.

Answer (Prob. 13.25) — Simply divide by the final coordinate to yield the inhomogeneous coordinates $(27/3, 18/3) = (9, 6)$.

Prob. 13.26 — Apply the following homogeneous transformation to the point $(3, 7)$, converting the

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

result back to the Euclidean plane:

$$\begin{bmatrix} 8 & 2 & 4 \\ 3 & 9 & 1 \\ 6 & 7 & 2 \end{bmatrix}$$

Answer (Prob. 13.26) — The result is

$$\begin{bmatrix} 8 & 2 & 4 \\ 3 & 9 & 1 \\ 6 & 7 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix} = \begin{bmatrix} 42 \\ 73 \\ 69 \end{bmatrix}.$$

Dividing by the third coordinate yields the point in the Euclidean plane: $(42/69, 73/69) = (0.609, 1.058)$.

Prob. 13.27 — Use homogeneous coordinates to simplify the computation of

(a) The point at the intersection of the lines $7x + 3y - 6 = 0$ and $y = -2x + 16$

(b) The line joining the points $(6, 2)$ and $(1, 9)$

Answer (Prob. 13.27) — The answers are given by computing the determinant of 3×3 matrices:

(a)

$$\left| \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 7 & 3 & -6 \\ -2 & -1 & 16 \end{bmatrix} \right| = (3 \cdot 16 - 6 \cdot 1)\mathbf{i} + (6 \cdot 2 - 7 \cdot 16)\mathbf{j} + (-1 \cdot 7 + 2 \cdot 3)\mathbf{k} = 42\mathbf{i} - 100\mathbf{j} - \mathbf{k} = \begin{bmatrix} 42 \\ -100 \\ -1 \end{bmatrix}.$$

The intersection is therefore the point $(-42, 100)$.

(b)

$$\left| \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 6 & 2 & 1 \\ 1 & 9 & 1 \end{bmatrix} \right| = (2 \cdot 1 - 9 \cdot 1)\mathbf{i} + (1 \cdot 1 - 6 \cdot 1)\mathbf{j} + (6 \cdot 9 - 2 \cdot 1)\mathbf{k} = -7\mathbf{i} - 5\mathbf{j} + 52\mathbf{k} = \begin{bmatrix} -7 \\ -5 \\ 52 \end{bmatrix}.$$

The line is therefore $-7x - 5y + 52 = 0$.

Prob. 13.28 — Use homogeneous coordinates to compute the intersection of the lines $8x + 2y - 3 = 0$ and $y = -4x + 12$. What do you notice about the result?

Answer (Prob. 13.28) — The result is given by

$$\left| \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 8 & 2 & -3 \\ -4 & -1 & 12 \end{bmatrix} \right| = (24 - 3)\mathbf{i} + (12 - 96)\mathbf{j} + (-8 + 8)\mathbf{k} = 21\mathbf{i} - 84\mathbf{j} + 0\mathbf{k} = \begin{bmatrix} 21 \\ -84 \\ 0 \end{bmatrix} \propto \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix}.$$

The intersection of the lines is the projective point $(1, -4, 0)$, which is the point at infinity associated with the slope -4. Note that although the lines are parallel, their intersection is well defined in the context of projective geometry.

Prob. 13.29 — Given the ellipse $4x^2 + 3y^2 = 100$,

(a) Draw the ellipse

(b) Construct the 3×3 matrix $\mathbf{C}$ representing the conic in homogeneous coordinates

(c) Compute the dual conic $\mathbf{C}^*$

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

(d) Use $\mathbf{C}$ to compute three different points on the conic, and add these points to your drawing.

(e) Use $\mathbf{C}^*$ to compute three different tangent lines to the conic, and add these lines to your drawing.

Answer (Prob. 13.29) — The answers are as follows:

(a) The ellipse is axis-aligned and is slightly taller than it is wide. It passes through the points $(0, \pm\sqrt{33.333})$ and $(\pm 5, 0)$.

(b) The matrix $\mathbf{C}$ satisfies $\begin{bmatrix} x & y & 1 \end{bmatrix} \mathbf{C} \begin{bmatrix} x & y & 1 \end{bmatrix}^\mathsf{T} = 0$, or

$$\mathbf{C} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -100 \end{bmatrix}$$

(c) The dual conic is

$$\mathbf{C}^* = \mathbf{C}^{-1} = \begin{bmatrix} 0.25 & 0 & 0 \\ 0 & 0.333 & 0 \\ 0 & 0 & -0.01 \end{bmatrix}.$$

(d) Answers may vary. One pair of points, for example, is given by $(1, \pm\sqrt{32})$, which is obtained by setting $x = 1$ and solving the above equation for y.

(e) Answers may vary. One pair of lines, for example, is given by $x + y = \pm 7.63$, which is obtained by setting $a = b$ and $c = 1$ and solving the equation $0.25a^2 + 0.333b^2 - 0.01c = 0$ for a. This leads to $0.583a^2 = 0.01$, or $a = b = \pm\sqrt{0.01/0.583} = 0.131$. The pair of lines is then given by $ax + by + c = 0$, or $0.131x + 0.131y = \pm 1$. Dividing by 0.131 yields $x + y = \pm 7.63$.

Prob. 13.30 — Which of the following statements is true?

(a) A similarity transformation is always an affine transformation.

(b) A projective transformation is always a similarity transformation.

(c) A similarity transformation is a Euclidean transformation if the scaling is 1.

Answer (Prob. 13.30) — (a) True. (b) False. (c) True.

Prob. 13.31 — Apply the following similarity transformation to the absolute points:

$$\begin{bmatrix} 5.196 & -3.000 & 6.789 \\ 3.000 & 5.196 & -8.312 \\ 0 & 0 & 1 \end{bmatrix}$$

Answer (Prob. 13.31) — The answer is

$$\begin{bmatrix} 5.196 & -3.000 & 6.789 \\ 3.000 & 5.196 & -8.312 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \pm j \\ 0 \end{bmatrix} = \begin{bmatrix} 5.196 \mp 3.000j \\ 3.000 \pm 5.196j \\ 0 \end{bmatrix}$$

Since homogeneous coordinates are not affected by scaling, we can multiply the result by $5.196 \pm 3.000j$ to yield

$$\begin{bmatrix} (5.196 \mp 3.000j)(5.196 \pm 3.000j) \\ (3.000 \pm 5.196j)(5.196 \pm 3.000j) \\ 0 \end{bmatrix} = \begin{bmatrix} 5.196^2 + 3.000^2 \\ \pm j(5.196^2 + 3.000^2) \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \pm j \\ 0 \end{bmatrix}$$

which is the expected result, since we know that the absolute points are invariant to similarity transformations.

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Prob. 13.32 — How many parameters specify the imaging process of a single pinhole camera? List these parameters, and specify for each whether it is extrinsic or intrinsic.

Answer (Prob. 13.32) — 11 parameters specify the imaging process of a single pinhole camera. Of these, 6 are extrinsic (3 for rotation, 3 for translation), and 5 are intrinsic (2 for principal point, 1 for focal length, 1 for aspect ratio, and 1 for skew).

Prob. 13.33 — List the two most common types of lens distortion.

Answer (Prob. 13.33) — The two most common types of lens distortion are radial and tangential.

Prob. 13.34 — Apply the normalized DLT algorithm to the following set of corresponding point pairs:

$\mathbf{x}$	(45,45)	(340,45)	(45,125)	(340,125)	(45,205)	(340,205)
$\mathbf{x}'$	(99,43)	(287,111)	(98,128)	(311,184)	(96,203)	(333,241)

Answer (Prob. 13.34) — The centroids are $(x_0, y_0) = (192.5, 125.0)$ and $(x_0', y_0') = (204.0, 151.7)$. The average distances are $d_{avg} = 161.0$ and $d'_{avg} = 123.4$. The scaling factors are $s = 0.0088$ and $s' = 0.0115$. The similarity transforms are therefore

$$\begin{aligned} T &= \begin{bmatrix}0.0088 & 0 & -1.6906; 0 & 0.0088 & -1.0978; 0 & 0 & 1\end{bmatrix} \\ T' &= \begin{bmatrix}0.0115 & 0 & -2.3379; 0 & 0.0115 & -1.7381; 0 & 0 & 1\end{bmatrix}\end{aligned}$$

The normalized points are $\{(-1.295, -0.703), (1.295, -0.703), (-1.295, 0), (1.295, 0), (-1.295, 0.703), (1.295, 0.703)\}$ and $\{(1.203, -1.245), (0.951, -0.466), (-1.215, -0.271), (1.226, 0.371), (-1.238, 0.588), (1.478, 1.024)\}$. From these the following matrix is constructed:

$$\breve{\mathbf{A}} = \begin{bmatrix} 1.2954 & 0.7026 & -1.0000 & 0 & 0 & 0 & 1.5587 & 0.8454 & -1.2033 \\ 0 & 0 & 0 & 1.2954 & 0.7026 & -1.0000 & 1.6132 & 0.8749 & -1.2453 \\ -1.2954 & 0.7026 & -1.0000 & 0 & 0 & 0 & 1.2322 & -0.6683 & 0.9512 \\ 0 & 0 & 0 & -1.2954 & 0.7026 & -1.0000 & -0.6037 & 0.3274 & -0.4660 \\ 1.2954 & 0 & -1.0000 & 0 & 0 & 0 & 1.5736 & 0 & -1.2148 \\ 0 & 0 & 0 & 1.2954 & 0 & -1.0000 & 0.3513 & 0 & -0.2712 \\ -1.2954 & 0 & -1.0000 & 0 & 0 & 0 & 1.5884 & 0 & 1.2262 \\ 0 & 0 & 0 & -1.2954 & 0 & -1.0000 & 0.4800 & 0 & 0.3705 \\ 1.2954 & -0.7026 & -1.0000 & 0 & 0 & 0 & 1.6033 & -0.8696 & -1.2377 \\ 0 & 0 & 0 & 1.2954 & -0.7026 & -1.0000 & -0.7621 & 0.4133 & 0.5883 \\ -1.2954 & -0.7026 & -1.0000 & 0 & 0 & 0 & 1.9150 & 1.0387 & 1.4784 \\ 0 & 0 & 0 & -1.2954 & -0.7026 & -1.0000 & 1.3262 & 0.7193 & 1.0238 \end{bmatrix}$$

Computing the SVD of $\breve{\mathbf{A}}$, taking the rightmost column of $\mathbf{V}$ and reshaping into a 3×3 matrix yields

$$\breve{\mathbf{H}} = \begin{bmatrix} 0.5060 & 0.1099 & 0.0821 \\ 0.1318 & 0.6372 & -0.0073 \\ 0.0547 & -0.0750 & 0.5413 \end{bmatrix}$$

The final normalized homography is

$$\mathbf{H} \leftarrow \mathbf{T}'^{-1}\breve{\mathbf{H}}\mathbf{T} = \begin{bmatrix} 0.4857 & -0.0500 & 30.3487 \\ 0.1738 & 0.3885 & -0.5504 \\ 0.0005 & -0.0007 & 0.5312 \end{bmatrix}$$

Applying this normalized homography to the first set of points yields

$\{(95.484, 47.311), (290.680, 114.358), (97.665, 118.658), (309.148, 174.966), (100.395, 207.986), (331.093, 246.983)\}$

which approximates the second set of points.

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Prob. 13.35 — Suppose you are given the following matrix representing the inverse of the image of the absolute conic (IAC). Compute the camera intrinsic parameters.

$$\boldsymbol{\omega}^{-1} = \mathbf{K}\mathbf{K}^{\mathsf{T}} = \begin{bmatrix} 231664.36 & 76044.66 & 335.00 \\ 76044.66 & 168561.41 & 227.00 \\ 335.00 & 227.00 & 1 \end{bmatrix}$$

Answer (Prob. 13.35) — From Equation (13.103), the internal calibration matrix is given by

$$\mathbf{K} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore,

$$\boldsymbol{\omega}^{-1} = \mathbf{K}\mathbf{K}^{\mathsf{T}} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha & 0 & 0 \\ \gamma & \beta & 0 \\ u_0 & v_0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha^2+\gamma^2+u_0^2 & \beta\gamma+u_0v_0 & u_0 \\ \beta\gamma+u_0v_0 & \beta^2+v_0^2 & v_0 \\ u_0 & v_0 & 1 \end{bmatrix}.$$

By inspection, then,

$$\begin{aligned} u_0 &= 335.0 \text{ pixels} \\ v_0 &= 227.0 \text{ pixels} \\ \beta &= \sqrt{168561.41 - v_0^2} = \sqrt{117032.41} = 342.1 \\ \gamma &= \frac{1}{\beta}(76044.66 - u_0v_0) = (-0.34/342.1) = -0.001 \\ \alpha &= \sqrt{231664.36 - \gamma^2 - u_0^2} = \sqrt{119439.36} = 345.6 \end{aligned}$$

From the equations just before Equation (13.104),

$$\begin{aligned} f_x &= \alpha = 345.6 \text{ pixels} \\ \theta &= \tan^{-1}(-f_x/\gamma) = \tan^{-1}(-345.6/0.001) = 1.57079 \text{ radians} \\ f_y &= \beta\sin\theta = (342.1)\sin(1.57079) = 342.1 \text{ pixels} \\ \text{aspect ratio} &= f_y/f_x = 342.1/345.6 = 0.99. \end{aligned}$$

Note: If, instead of 76044.66, the second and fourth elements of the matrix are 76045.34, then all answers will be the same except that $\gamma = 0.001$ instead of $\gamma = -0.001$. But this positive value of γ leads to a physically implausible solution, with $\theta = -1.57079$ radians and $f_y = -342.1$ pixels.

Prob. 13.36 — Calibrate an actual camera using Zhang's algorithm. That is, obtain access to a camera, download an implementation of a camera calibration routine (nearly all implementations are some variation of Zhang's algorithm), print out a chessboard pattern and tape it to a hard flat surface, capture several images, obtain coordinates of the corners either manually or automatically, and feed the coordinates into the code to compute the intrinsic parameters. (Ignore the extrinsic parameters.)

Answer (Prob. 13.36) — Answers may vary.

Prob. 13.37 — Define the following terms: epipole, epipolar line, epipolar plane, and epipolar constraint.

Answer (Prob. 13.37) — The answers are as follows:

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

- Epipole: the intersection of the baseline with the image plane.
- Epipolar line: the intersection of the epipolar plane with the image plane.
- Epipolar plane: the plane defined by the two centers of projection and a scene point.
- Epipolar constraint: the geometric fact that corresponding points in a stereo pair of images must lie on their respective epipolar lines.

Prob. 13.38 — Explain the relationship between the fundamental matrix and the essential matrix.

Answer (Prob. 13.38) — The fundamental matrix captures the geometric relationship between two uncalibrated cameras, whereas the essential matrix captures the geometric relationship between two calibrated cameras. if the cameras are calibrated, the fundamental matrix can be converted to the essential matrix, and vice versa.

Prob. 13.39 — Prove the following additional properties of essential matrices. For simplicity, just prove the "only if" part, that is, show that these properties hold for any essential matrix:

(a) A real 3×3 matrix $\mathbf{E}$ is an essential matrix if and only if $\det(\mathbf{E}) = 0$ and $\frac{1}{2} tr^2(\mathbf{E}\mathbf{E}^\mathsf{T}) = tr((\mathbf{E}\mathbf{E}^\mathsf{T})^2)$.

(b) A real 3×3 matrix $\mathbf{E}$ is an essential matrix if and only if $\mathbf{E}\mathbf{E}^\mathsf{T}\mathbf{E} = \frac{1}{2} tr(\mathbf{E}\mathbf{E}^\mathsf{T})\mathbf{E}$.

Answer (Prob. 13.39) — (a) The "only if" part follows from simple algebraic manipulation:

$$\mathbf{E}\mathbf{E}^\mathsf{T} = [\mathbf{t}]_\times [\mathbf{t}]_\times^\mathsf{T} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} 0 & t_z & -t_y \\ -t_z & 0 & t_x \\ t_y & -t_x & 0 \end{bmatrix} = \begin{bmatrix} t_y^2 + t_z^2 & -t_x t_y & -t_x t_z \\ -t_x t_y & t_x^2 + t_z^2 & -t_y t_z \\ -t_x t_z & -t_y t_z & t_x^2 + t_y^2 \end{bmatrix} \tag{13.204}$$

$$(\mathbf{E}\mathbf{E}^\mathsf{T})^2 = \begin{bmatrix} (t_y^2 + t_z^2)^2 + t_x^2 t_y^2 + t_x^2 t_z^2 & \cdot & \cdot \\ \cdot & (t_x^2 + t_z^2)^2 + t_x^2 t_y^2 + t_y^2 t + z^2 & \cdot \\ \cdot & \cdot & (t_x^2 + t_y^2)^2 + t_x^2 t_z^2 + t_y^2 t_z^2 \end{bmatrix}, \tag{13.205}$$

where we have used $\cdot$ to indicate the elements that we do not care about. From these equations it is clear that $tr(\mathbf{E}\mathbf{E}^\mathsf{T}) = 2(t_x^2 + t_y^2 + t_z^2)$, so that $\frac{1}{2} tr^2(\mathbf{E}\mathbf{E}^\mathsf{T}) = 2(t_x^2 + t_y^2 + t_z^2)^2$, which is also obtained by adding the diagonal elements of the second matrix.

(b) Again, the "only if" part is easy to show:

$$\begin{aligned}
\mathbf{E}\mathbf{E}^\mathsf{T}\mathbf{E} &= \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} 0 & t_z & -t_y \\ -t_z & 0 & t_x \\ t_y & -t_x & 0 \end{bmatrix} \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \mathbf{R} && (13.206) \\
&= \begin{bmatrix} 0 & -t_y^2 t_z - t_z^3 - t_x^2 t_z & t_y^3 + t_y t_z^2 + t_x^2 t_y \\ t_x^2 t_z + t_z^3 + t_y^2 t_z & 0 & -t_x t_y^2 - t_x^3 - t_x t_z^2 \\ -t_y t_z^2 - t_x^2 t_y - t_y^3 & t_x t_z^2 + t_x^3 + t_x t_y^2 & 0 \end{bmatrix} \mathbf{R} && (13.207) \\
&= (t_x^2 + t_y^2 + t_z^2) \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \mathbf{R} && (13.208) \\
&= \frac{1}{2} tr(\mathbf{E}\mathbf{E}^\mathsf{T})\mathbf{E}. && (13.209)
\end{aligned}$$

Prob. 13.40 — Prove that the essential matrix is given by $\mathbf{E} = \mathbf{R}\left[\mathbf{R}^\mathsf{T}\mathbf{t}\right]_\times$, as stated in Equation (13.182).

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Answer (Prob. 13.40) — From epipolar geometry we have

$$\begin{aligned} 0 &= \bar{\mathbf{x}}'^{\mathsf{T}}(\mathbf{t} \times \mathbf{R}\bar{\mathbf{x}}) \\ &= \bar{\mathbf{x}}'^{\mathsf{T}}(\mathbf{R}\mathbf{R}^{\mathsf{T}}\mathbf{t} \times \mathbf{R}\bar{\mathbf{x}}) \\ &= \bar{\mathbf{x}}'^{\mathsf{T}}\mathbf{R}(\mathbf{R}^{\mathsf{T}}\mathbf{t} \times \bar{\mathbf{x}}) \\ &= \bar{\mathbf{x}}'^{\mathsf{T}} \underbrace{\left(\mathbf{R}\left[\mathbf{R}^{\mathsf{T}}\mathbf{t}\right]_{\times}\right)}_{\mathbf{E}} \bar{\mathbf{x}}. \end{aligned}$$

Prob. 13.41 — (a) Estimate the fundamental matrix from the following set of correspondences, (b) compute the epipolar line for each point, and (c) display the epipolar lines on a pair of plots:

$\mathbf{x}$	(190,155)	(420,114)	(252,29)	(150,111)	(35,228)	(443,230)	(149,240)	(276,324)
$\mathbf{x}'$	(231,138)	(442,98)	(272,9)	(169,91)	(58,209)	(460,217)	(184,225)	(312,310)

Answer (Prob. 13.41) — (a) The fundamental matrix is computed using Algorithm 13.13. First the similarity transforms are computed using Equation (13.118):

$$(x_0, y_0) = (239.4, 178.9) \tag{13.210}$$
$$(x_0', y_0') = (266.0, 162.1) \tag{13.211}$$
$$d_{avg} = 55.6 \tag{13.212}$$
$$d_{avg}' = 55.3 \tag{13.213}$$
$$s = 0.0254 \tag{13.214}$$
$$s' = 0.0256 \tag{13.215}$$
$$\mathbf{T} = \begin{bmatrix} 0.0254 & 0 & -6.092 \\ 0 & 0.0254 & -4.552 \\ 0 & 0 & 1 \end{bmatrix} \tag{13.216}$$
$$\mathbf{T}' = \begin{bmatrix} 0.0256 & 0 & -6.804 \\ 0 & 0.0256 & -4.147 \\ 0 & 0 & 1 \end{bmatrix} \tag{13.217}$$

The normalized coordinates are computed as $\breve{\mathbf{x}} = \mathbf{T}\mathbf{x}$ and $\breve{\mathbf{x}}' = \mathbf{T}'\mathbf{x}'$:

$\breve{\mathbf{x}}$	$\breve{\mathbf{x}}'$
(-1.2566, -0.6076)	(-0.8953, -0.6171)
(4.5969, -1.6511)	(4.5021, -1.6403)
(0.3213, -3.8143)	(0.1535, -3.9170)
(-2.2746, -1.7274)	(-2.4813, -1.8194)
(-5.2013, 1.2502)	(-5.3207, 1.1991)
(5.1822, 1.3011)	(4.9626, 1.4037)
(-2.3000, 1.5556)	(-2.0976, 1.6084)
(0.9321, 3.6934)	(1.1767, 3.7827)

The matrix $\breve{\mathbf{A}}$ is created by plugging the normalized coordinates into Equation (13.187). Then $\breve{\mathbf{A}}\breve{\mathbf{f}} = \mathbf{0}$ is solved for $\breve{\mathbf{f}}$. (This can be done by computing the SVD of $\breve{\mathbf{A}}$, then taking the rightmost right singular vector.)

$$\breve{\mathbf{f}} = \begin{bmatrix} -0.0104 & -0.1341 & -0.2601 & 0.1223 & -0.0002 & -0.6155 & 0.2748 & 0.6560 & 0.1211 \end{bmatrix}^{\mathsf{T}}.$$

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Rearranging $\breve{\mathbf{f}}$ yields the matrix

$$\breve{\mathbf{F}} = \begin{bmatrix} -0.0104 & -0.1341 & -0.2601 \\ 0.1223 & -0.0002 & -0.6155 \\ 0.2748 & 0.6560 & 0.1211 \end{bmatrix}$$

The SVD of $\breve{\mathbf{F}}$ is given by

$$\breve{\mathbf{U}} = \begin{bmatrix} -0.3393 & -0.2076 & 0.9175 \\ -0.4429 & -0.8252 & -0.3505 \\ 0.8299 & -0.5253 & 0.1880 \end{bmatrix} \qquad \breve{\mathbf{\Sigma}} = \begin{bmatrix} 0.7697 & 0 & 0 \\ 0 & 0.6385 & 0 \\ 0 & 0 & 0.0008 \end{bmatrix} \qquad \breve{\mathbf{V}} = \begin{bmatrix} 0.2305 & -0.3807 & -0.8955 \\ 0.7665 & -0.4958 & 0.4081 \\ 0.5994 & 0.7805 & -0.1775 \end{bmatrix}$$

Setting the smaller singular value to zero yields a rank-2 matrix:

$$\begin{aligned} \breve{\mathbf{F}} &= \breve{\mathbf{U}}\, diag(0.7697, 0.6385, 0)\breve{\mathbf{V}}^{\mathsf{T}} \\ &= \begin{bmatrix} -0.3393 & -0.2076 & 0.9175 \\ -0.4429 & -0.8252 & -0.3505 \\ 0.8299 & -0.5253 & 0.1880 \end{bmatrix} \begin{bmatrix} 0.7697 & 0 & 0 \\ 0 & 0.6385 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.2305 & -0.3807 & -0.8955 \\ 0.7665 & -0.4958 & 0.4081 \\ 0.5994 & 0.7805 & -0.1775 \end{bmatrix}^{\mathsf{T}} \\ &= \begin{bmatrix} -0.0097 & -0.1344 & -0.2600 \\ 0.1220 & -0.0001 & -0.6156 \\ 0.2749 & 0.6559 & 0.1211 \end{bmatrix} \end{aligned}$$

Unnormalizing the result yields the desired fundamental matrix:

$$\begin{aligned} \mathbf{F} &= \mathbf{T}'^{\mathsf{T}}\breve{\mathbf{F}}\mathbf{T} \\ &= \begin{bmatrix} 0.0256 & 0 & -6.804 \\ 0 & 0.0256 & -4.147 \\ 0 & 0 & 1 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} -0.0097 & -0.1344 & -0.2600 \\ 0.1220 & -0.0001 & -0.6156 \\ 0.2749 & 0.6559 & 0.1211 \end{bmatrix} \begin{bmatrix} 0.0254 & 0 & -6.092 \\ 0 & 0.0254 & -4.552 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -0.0000 & -0.0001 & 0.0105 \\ 0.0001 & -0.0000 & -0.0348 \\ -0.0042 & 0.0400 & -1.7042 \end{bmatrix} \end{aligned}$$

(b) The epipolar line associated with the first point of the first set is given by

$$\mathbf{x}'^{\mathsf{T}}\mathbf{F}\mathbf{x} = \mathbf{x}'^{\mathsf{T}} \begin{bmatrix} -0.0000 & -0.0001 & 0.0105 \\ 0.0001 & -0.0000 & -0.0348 \\ -0.0042 & 0.0400 & -1.7042 \end{bmatrix} \begin{bmatrix} 190 \\ 155 \\ 1 \end{bmatrix} = \mathbf{x}'^{\mathsf{T}} \begin{bmatrix} -0.0042 \\ -0.0197 \\ 3.6954 \end{bmatrix}$$

The epipolar line in the second image associated with this point is $-0.0042x' - 0.0197y' + 3.6954 = 0$. The epipolar line associated with the first point of the second set is given by

$$\mathbf{x}'^{\mathsf{T}}\mathbf{F}\mathbf{x} = \begin{bmatrix} 231 & 138 & 1 \end{bmatrix} \begin{bmatrix} -0.0000 & -0.0001 & 0.0105 \\ 0.0001 & -0.0000 & -0.0348 \\ -0.0042 & 0.0400 & -1.7042 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0.0053 & 0.0198 & -4.0695 \end{bmatrix} \mathbf{x}$$

The epipolar line in the first image associated with this point is $0.0053x + 0.0198y - 4.0695 = 0$. The other epipolar lines are determined in a similar manner.

(c) Plotting the epipolar lines is straightforward.

Prob. 13.42 — Decompose the following essential matrix into rotation and translation.

$$\begin{bmatrix} -0.8497 & -0.9437 & 9.0579 \\ 0.9437 & -0.8497 & -8.1472 \\ -1.2798 & 12.1155 & 0 \end{bmatrix}$$

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Answer (Prob. 13.42) — The essential matrix is decomposed as follows:

$$\mathbf{R} = \begin{bmatrix} 0.7431 & -0.66910 & \\ 0.6691 & 0.7431 & 0 \\ 0 & 0 & 1.0000 \end{bmatrix} \qquad \mathbf{t} = \begin{bmatrix} 8.1472 \\ 9.0579 \\ 1.2699 \end{bmatrix}$$

Prob. 13.43 — Implement block-based matching of a pair of rectified stereo images, using the SAD dissimilarity measure. For efficiency, your code should precompute the 3D array of dissimilarities, followed by a series of separable convolutions (one pair of convolutions per disparity). Implement the left-to-right consistency check, retaining a value in the left disparity map only if the corresponding point in the right disparity map agrees in its disparity. The resulting disparity map should be valid only at the pixels that pass the consistency check; set other pixels to zero. (*Note*: For simplicity, do not worry about setting the values of pixels along the left border of the left image.)

Answer (Prob. 13.43) — Answers may vary.

Prob. 13.44 — Implement the detection and tracking of sparse features points throughout a video sequence. To detect good features in the first frame, use either the Harris corner detector or the Tomasi-Kanade method, as explained in Chapter 7. Then, for each pair of consecutive frames, perform Lucas-Kanade tracking of all the features to update their 2D image positions. Remember to keep the feature coordinates as floating point values throughout the tracking process, only rounding for display purposes; to handle noninteger values, use bilinear interpolation. Do not worry about declaring features lost, but simply allow them to continue tracking throughout the sequence, even if they drift to a neighboring surface in the world due to occlusion. Nevertheless, be sure to perform bounds checking so that features that reach the image border do not cause the program to crash due to out-of-bounds memory access.

Answer (Prob. 13.44) — Answers may vary.

© 2018 Cengage Learning®. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.