Exam

Computer Vision

***Human Vision and Computer Vision* course**

Master CIMET - June 2011

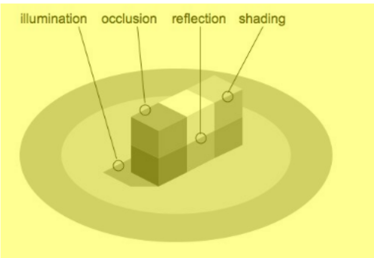
(All documents authorized - duration: 2 hours)

Problem 1: Short Answer Questions

This part consists of four short answer questions. Answer each question in this part clearly and concisely. Marks will be deducted for overly long or unclear answers.

Question 1: Name four scene properties that would cause an edge (brightness discontinuity) in an image.

_ a depth discontinuity (i.e., a foreground/background segmentation)
_ a surface orientation discontinuity (e.g., two intersecting planar surfaces)
_ a reflectance discontinuity (i.e., a change in surface colour/material on an otherwise smooth surface)
_ illumination boundaries (e.g., cast shadows, light sources, specularities)



Question 2: Compare and contrast the use of stereopsis and motion information in vision systems. In particular you should explain what is meant by the terms stereo disparity and optic flow, and consider:

(a) differences between the algorithms used for estimating stereo disparity in pairs of images and optic flow fields in image sequences;

(b) differences between the information for layout contained in stereo disparities and optic flow vectors.

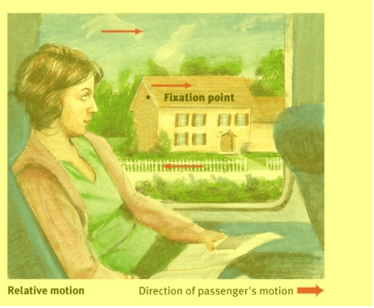
See Lecture 4

Question 3: Consider a train moving on straight tracks over flat, open countryside. Describe, with a sketch, the optic flow field observed by a camera aimed horizontally out of a window on the left side of the train, assuming the camera is on a support fixed rigidly to the train. In addition, describe the flow field produced if the camera is instead mounted on an active head, and is moved so as to track an object on the ground to the side of the train.

See Lecture 2 Part1

(a)Camera inside the train

If the observer moves through a stationary environment, the resulting movement is called motion parallax. Objects will move at different speeds on your retina (for a particular speed of observer movement and choice of fixation object) depending on their distance from the observer.

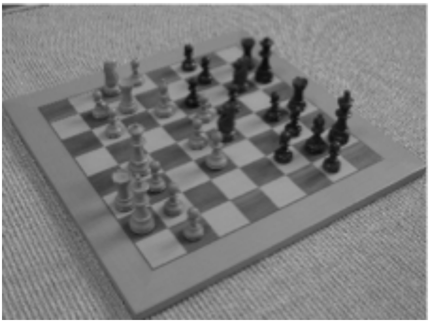


Smaller and closer objects move faster, far away objects and big ones move slowly. The optical flow field depend on the distance of the objects to the camera. It decreases with depth.

(b) camera outside the train

Far away objects and the one being tracked will move along the direction of the train. The rest will move in the opposite direction. The optical flow increases with depth.

Question 4: What are the main sources of information for layout (the three-dimensional arrangement of surfaces) in each of the images below? Relate your answer explicitly to features of the two images, and state any assumptions that need to be made in order to use each information source.



Blur, Texture gradient, Occlusion

Perspective, 2 vanishing points, size of squares and of pieces

Problem 2: Multiple TRUE/FALSE Questions

Each question in this part requires you to indicate which of four statements, (A)–(D), are **true** and which are **false**? Note that questions in this part may have zero, one or more statements that are true. No statement is intended to be ambiguous. You may add a comment to explain your choice of **true** or **false** if you find a statement to be ambiguous.

Question 1: Which of the following statements are **true** of a pinhole camera? Which are **false**?

(A) A pinhole camera is a box with a small hole in it.

(B) Images in a pinhole camera are upside down.

(C) A pinhole camera has a fixed focal length, f.

(D) Images in a pinhole camera are a perspective projection.

Solution : (A) True, (B) True, (C) False, (D) True.

Question 2: A pinhole camera is not widely used in practice. Which of the following statements are **true** of an actual pinhole camera? Which are **false**?

(A) It takes too long to acquire an image.

(B) It uses an orthographic projection.

(C) It works only for black and white (B&W), not for color.

(D) It has too small a depth of field (i.e., too small a range of object distances for which the image is in sharp focus).

Solution : (A) True, (B) False, (C) False, (D) False.

Question 3: Consider a set of 3D lines that all share the same horizon line when projected onto an image under a perspective projection. Which of the statements, (A)–(D), are **true** and which are **false**?

(A) The lines are all parallel in 3D.

(B) The lines all lie on the same 3D plane.

(C) The lines are both parallel and on the same 3D plane.

(D) There is no constraint. A set of 3D lines always shares the same horizon line when projected onto an image?

Solution : (A) False, (B) True, (C) False, (D) False.

Question 4: Consider conditions under which an epipolar constraint used in stereo matching holds between images from two cameras. Which of the following conditions are **true**? Which are **false**? Note: You can assume that the cameras perform standard perspective projection.

(A) The two cameras must have coplanar projection planes.

(B) The two cameras must face in the same direction (i.e., have parallel optical axes).

(C) The two images must be rectified.

(D) There are no restrictions on camera locations or orientations, an epipolar constraint always applies.

Solution : (A) False, (B) False, (C) False, (D) True.

Question 5: Stereo matching can be performed by correlating windows of pixels between the two images. But, it is difficult to know what window size to use. The following statements identify problems when the selected window size is too large. Which are **true**? Which are **false**?

(A) There will be more false matches due to ambiguity and image noise.

(B) The exact location of correct matches will be known with less accuracy.

(C) Places where depth is discontinuous will be poorly matched.

(D) The epipolar constraint is not as effective to limit the number of matches.

Solution : (A) False, (B) True, (C) True, (D) False.

Problem 3:

Suppose that a camera is at an initial position in which its coordinate system coincides with the coordinate system of the scene, so that the camera matrix can be expressed as $\mathbf{P} = \mathbf{K}\,[\mathbf{I} \mid \mathbf{0}]$, where $\mathbf{K}$ is the calibration matrix. The camera takes a view of the scene, then undergoes a pure translation $\mathbf{T}$, and then takes a second view. There is no change in the calibration between the two views and there is no camera rotation during the displacement.

Question 1: Assuming that $\mathbf{T}$ is expressed in the coordinate system of that camera, what is the form of the camera matrix after translation in relation to $\mathbf{K}$ and $\mathbf{T}$? (Let us note $\mathbf{P'}$ this matrix)

(a) The general camera matrix is $\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I} \mid -\tilde{\mathbf{C}}]$ where $\tilde{\mathbf{C}}$ is the position of the camera center in scene coordinates. Here the scene coordinates are the camera coordinates for the first camera position. $\mathbf{T}$ is the translation of the camera center from its original position to its second position in scene coordinates. Therefore, $\mathbf{T} = \mathbf{CC'} = \tilde{\mathbf{C}}$.
For pure translation, the rotation $\mathbf{R}$ is the identity matrix. Therefore the camera matrix in this case is $\mathbf{P'} = \mathbf{K}\,[\mathbf{I} \mid -\mathbf{T}]$.

Question 2: Show that the fundamental matrix for the two images depends only on the epipole $\mathbf{e'}$ of the second camera. (Let us note $\mathbf{F}$ □□$[\mathbf{e'}]_{\mathrm{x}}\, \mathbf{P'P}$□ the fundamental matrix and $\mathbf{P}$□ the pseudo-inverse of $\mathbf{P}$, i.e. $\mathbf{PP}$□$= \mathbf{I}$).

One form of the fundamental matrix is $\mathbf{F} = [\mathbf{e'}]_{\mathrm{X}}\, \mathbf{P'P}^{+}$. For this problem, we have $\mathbf{P} = \mathbf{K}\,[\mathbf{I} \mid \mathbf{0}]$ and $\mathbf{P'} = \mathbf{K}\,[\mathbf{I} \mid -\mathbf{T}]$. First, we calculate

We have. $\mathbf{P} = [\mathbf{K} \mid \mathbf{0}]$, therefore $\mathbf{P}^{\mathbf{T}} = \begin{bmatrix} \mathbf{K}^{\mathbf{T}} \\ 0_3{}^{\mathbf{T}} \end{bmatrix}$ and $\mathbf{P}\mathbf{P}^{\mathbf{T}} = \mathbf{K}\mathbf{K}^{\mathbf{T}}$, so that $(\mathbf{P}\mathbf{P}^{\mathbf{T}})^{-1} = \mathbf{K}^{-\mathbf{T}}\mathbf{K}^{-1}$

Therefore $\mathbf{P}^{+} = \begin{bmatrix} \mathbf{K}^{\mathbf{T}} \\ \mathbf{0}_3^{\mathbf{T}} \end{bmatrix} \mathbf{K}^{-\mathbf{T}}\mathbf{K}^{-1} = \begin{bmatrix} \mathbf{K}^{\mathbf{T}}\mathbf{K}^{-\mathbf{T}}\mathbf{K}^{-1} \\ \mathbf{0}_3^{\mathbf{T}} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{-1} \\ \mathbf{0}_3^{\mathbf{T}} \end{bmatrix}$ and $\mathbf{P'}\mathbf{P}^{+} = [\mathbf{K} \mid -\mathbf{K}\,\mathbf{T}] \begin{bmatrix} \mathbf{K}^{-1} \\ \mathbf{0}_3^{\mathbf{T}} \end{bmatrix} = \mathbf{I}$

Going back to the expression of $\mathbf{F}$, this expression is reduced to $\mathbf{F} = [\mathbf{e'}]_{\mathrm{X}}$

Question 3: An image point $\mathbf{x}$ in the initial view moves to a point $\mathbf{x'}$ because of the translation. The point $\mathbf{x'}$ is on the epipolar line of $\mathbf{x}$, which passes through the epipole $\mathbf{e'}$. (Let us note $\mathbf{l'} = [\mathbf{e'}]_{\mathrm{x}}\, \mathbf{x}$ the epipolar line of $\mathbf{x}$). Show that the point $\mathbf{x}$ also belongs to that epipolar line, and therefore that for pure translation displacements, points move along epipolar lines. (Let us note that the epipole in the case of pure translation does not move from view to view and corresponds to the focus of expansion of optical flow).

(c) The epipolar line of $\mathbf{x}$ is $\mathbf{l'} = [\mathbf{e'}]_\times \mathbf{x}$. We want to show that $\mathbf{x}$ is on $\mathbf{l'}$, i.e that $\mathbf{x}^T \mathbf{l'}=0$. This is the case because $\mathbf{x}^T [\mathbf{e'}]_\times \mathbf{x} = \mathbf{0}$, a property resulting from the fact that $[\mathbf{e'}]_\times$ is skew symmetric (Fundamental matrix, slide 10).

Question 4: The translation is now assumed parallel to the *x*-axis of the coordinate system of the camera in its initial position. What form does the fundamental matrix take in this special case?

(d) If the camera translation is parallel to the x-axis, then $\mathbf{e'} = (1, 0, 0)^T$. Then $\mathbf{F} = [\mathbf{e'}]_\times$ becomes

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Problem 4:

The images shown in the two figures (a) and (b) below are line drawings from the left and right parts respectively of a stereo pair, taken using identical cameras mounted at the same horizontal level and with their optical axes parallel.

Question 1: Draw a plan view of the scene to show roughly what the spatial arrangement of the objects was. (Relative rather than accurately measured positions only are required.) Describe the reasoning behind your answer.

See below

Question 2: How could these images be used to make estimates of absolute distances to the objects in the scene? What additional information about the cameras is required for this?

The camera position to avoid ambiguity. Intrinsic and extrinsic parameters are requested to compute (depth) disparity.

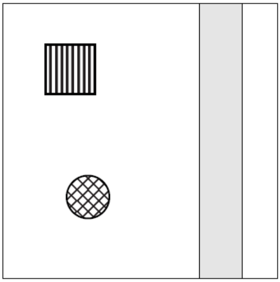
$X_{left}-X_{right} = B\ f/Z$

With B= baseline distance, F= focal, Z= absolute depth, $X_{left}-X_{right}$ = disparity.

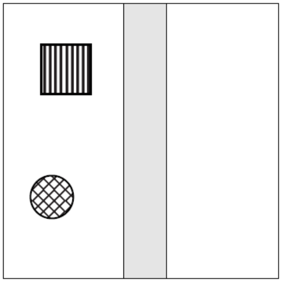
Question 3: Discuss the issues that arise when matching corresponding points in a stereo pair, with particular reference to the difference between a pair of the kind shown in the two figures (a) and (b) below and a pair of the kind shown in the two figures (c) and (d) below.

In the figures a and b the distance between corresponding points increases with depth, as X_{left}-X_{right} = B f/Z = disparity, this equation is inversely proportional to depth.

In the figures c and d it will be very difficult to match "interest points" between left image and right image, many confusion can happen.



(a)



(b)

Little shift from right to left

Bigger shift from right to left => depth is lower

Biggest shift from right to left => depth is lower

No vertical disparity

(c)

(d)