Computer Vision

Spring 2006 15-385,-685

Instructor: S. Narasimhan

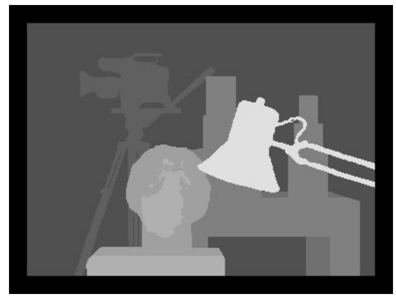
Wean 5403
T-R 3:00pm – 4:20pm

Lecture #15

Binocular Stereo - Calibration

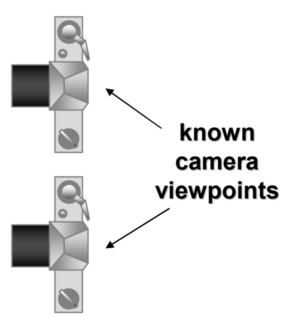
Lecture #15

Binocular Stereo

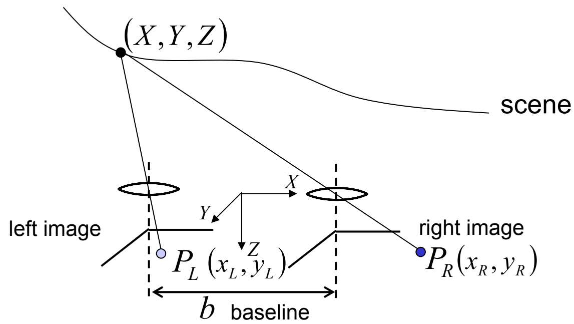


Stereo Reconstruction - RECAP

- The Stereo Problem
  - Shape from two (or more) images
  - Biological motivation



Disparity and Depth - RECAP

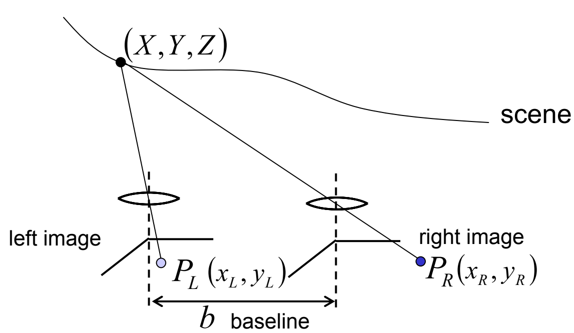


Assume that we know P_L corresponds to P_R

From perspective projection (define the coordinate system as shown above)

$$\frac{x_L}{f} = \frac{X + b/2}{Z} \qquad \frac{x_R}{f} = \frac{X - b/2}{Z} \qquad \frac{y_L}{f} = \frac{y_R}{f} = \frac{Y}{Z}$$

Disparity and Depth - RECAP



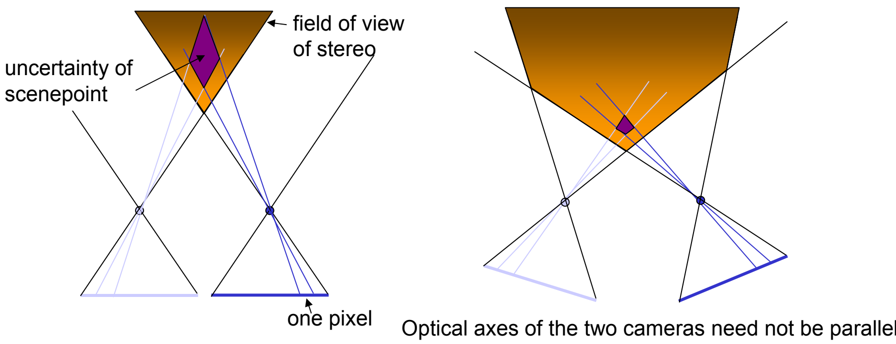
$$\frac{x_L}{f} = \frac{X + b/2}{Z} \qquad \frac{x_R}{f} = \frac{X - b/2}{Z} \qquad \frac{y_L}{f} = \frac{y_R}{f} = \frac{Y}{Z}$$

$$\Rightarrow \quad X = \frac{b(x_L + x_R)}{2(x_L - x_R)} \qquad Y = \frac{b(y_L + y_R)}{2(x_L - x_R)} \qquad Z = \frac{bf}{(x_L - x_R)}$$

$d = x_L - x_R$ is the **disparity** between corresponding left and right image points

- inverse proportional to depth
- disparity increases with baseline **b**

Vergence



- Field of view decreases with increase in baseline and vergence
- Accuracy increases with baseline and vergence

Binocular Stereo Calibration

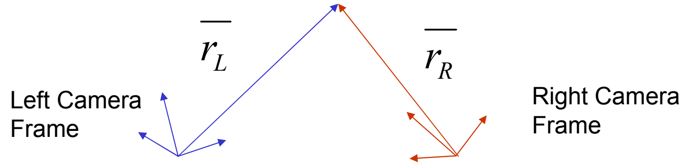
• RELATIVE ORIENTATION:

We need to know position and orientation of one camera with respect to the other, before computing depth of scene points.

• ABSOLUTE ORIENTATION:

We may need to know position and orientation of a stereo system with respect to some external system (for example, a 3D scanner).

Binocular Stereo Calibration - Notation



- We need to transform one coordinate frame to another.
- The transformation includes a rotation and a translation:

$$\overline{r}_R = R\,\overline{r}_L + \overline{r}_0$$

$\overline{r}_0$: Translation of Left frame w.r.t Right

R : Rotation of Left frame w.r.t Right

Binocular Stereo Calibration - Notation

- In matrix notation, we can write $\overline{r}_R = R\,\overline{r}_L + \overline{r}_0$ as:

$$\overline{r}_L = \begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix} \qquad \overline{r}_R = \begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix}$$

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \qquad \overline{r}_0 = \begin{bmatrix} r_{14} \\ r_{24} \\ r_{34} \end{bmatrix}$$

Binocular Stereo Calibration - Notation

- We can expand $\overline{r_R} = R\,\overline{r_L} + \overline{r_0}$ as:

$$\boxed{\begin{aligned} r_{11}\, x_L + r_{12}\, y_L + r_{13}\, z_L + r_{14} &= x_R \\ r_{21}\, x_L + r_{22}\, y_L + r_{23}\, z_L + r_{24} &= y_R \\ r_{31}\, x_L + r_{32}\, y_L + r_{33}\, z_L + r_{34} &= z_R \end{aligned}}$$

Orthonormality Constraints $R^T R = I$

(a) Rows of R are perpendicular vectors

$$\begin{aligned} r_{11}\, r_{21} + r_{12}\, r_{22} + r_{13}\, r_{23} &= 0 \\ r_{21}\, r_{31} + r_{22}\, r_{32} + r_{23}\, r_{33} &= 0 \\ r_{11}\, r_{31} + r_{12}\, r_{32} + r_{13}\, r_{33} &= 0 \end{aligned}$$

(b) Each row of R is a unit vector

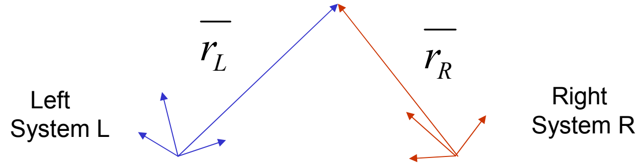
$$\begin{aligned} r_{11}^2 + r_{12}^2 + r_{13}^2 &= 1 \\ r_{21}^2 + r_{22}^2 + r_{23}^2 &= 1 \\ r_{31}^2 + r_{32}^2 + r_{33}^2 &= 1 \end{aligned}$$

NOTE: Constraints are NON-LINEAR!

They can only be used once (do not change with scene points).

Absolute Orientation

- We have measured some scene points using Stereo System L and Stereo System R
- Find Orientation (Translation and Rotation) of system L w.r.t system R.
- Useful for merging partial depth information from different views.



- Problem:

Given $\overline{r_L} = (x_L, y_L, z_L)$ $\overline{r_R} = (x_R, y_R, z_R)$

Find R $\overline{r_0}$ $(r_{11}, r_{12}, \ldots, r_{34})$

How many scene points are needed?

- Each scene point gives 3 equations:

$$r_{11}\, x_L + r_{12}\, y_L + r_{13}\, z_L + r_{14} = x_R$$
$$r_{21}\, x_L + r_{22}\, y_L + r_{23}\, z_L + r_{24} = y_R$$
$$r_{31}\, x_L + r_{32}\, y_L + r_{33}\, z_L + r_{34} = z_R$$

- Six additional equations from orthonormality of Rotation matrix
- For n scene points, we have (3n + 6) equations and 12 unknowns
- It appears that 2 scene points suffice. But orthonormal constraints are non-linear.

THREE NON-COLLINEAR SCENE POINTS ARE SUFFICIENT (see Horn)

Solving an Over-determined System

- Generally, more than 3 points are used to find the 12 unknowns

- Formulate Error for scene point i as:

$$e_i = (R\,\overline{r_{L_i}} + \overline{r_0}) - \overline{r_{R_i}}$$

- Find R & $\overline{r_0}$ that minimize:

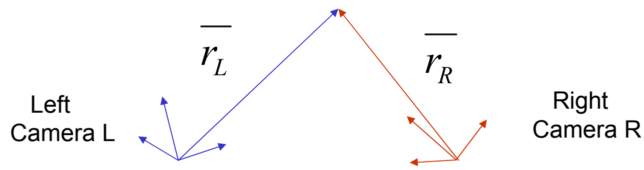
$$E = \sum_{i=1}^{N} |e_i|^2 \; + \; \lambda\,(R^T R - I)$$

Orthonormality Constraint

Relative Orientation

Relative Orientation

- Find Orientation (Translation and Rotation) between two cameras (within the same stereo system).
- We must do this before using the stereo system.



- Problem:

Here, we DO NOT know both $\overline{r}_L$ $\overline{r}_R$

We only know the image coordinates (x'_L, y'_L) (x'_R, y'_R)

in the two cameras CORRESPOND to the same scene point!

Relative Orientation

- Again, we start with: $\overline{r}_R = R\,\overline{r}_L + \overline{r}_0$
- OR:

$$r_{11}\, x_L + r_{12}\, y_L + r_{13}\, z_L + r_{14} = x_R$$
$$r_{21}\, x_L + r_{22}\, y_L + r_{23}\, z_L + r_{24} = y_R$$
$$r_{31}\, x_L + r_{32}\, y_L + r_{33}\, z_L + r_{34} = z_R$$

- Assume: we know focal length of both cameras
- Then: $\dfrac{x'_L}{f} = \dfrac{x_L}{z_L}$ & $\dfrac{y'_L}{f} = \dfrac{y_L}{z_L}$ (same for right camera)
- Hence,

$$r_{11}\, x'_L + r_{12}\, y'_L + r_{13}\, f + r_{14}(f / z_L) = x'_R\,(z_R / z_L)$$
$$r_{21}\, x'_L + r_{22}\, y'_L + r_{23}\, f + r_{24}(f / z_L) = y'_R\,(z_R / z_L)$$
$$r_{31}\, x'_L + r_{32}\, y'_L + r_{33}\, f + r_{34}(f / z_L) = f\,(z_R / z_L)$$

Problem Formulation

We know: $(x'_L, y'_L)\ (x'_R, y'_R)\ \&\ \ f$

Find: $(r_{11}, r_{12}, \ldots, r_{34})\ \ (z_L, z_R)$

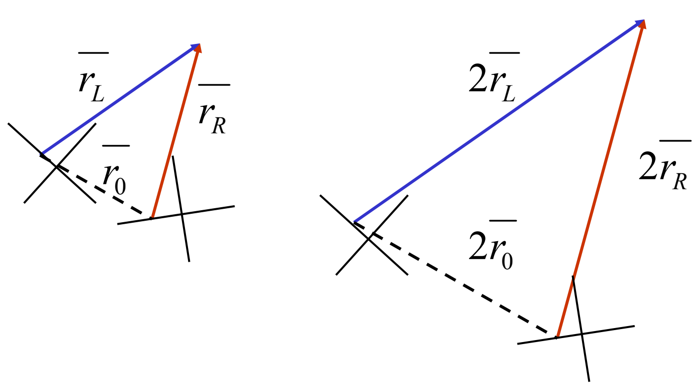
That satisfies :

$$
\begin{aligned}
r_{11}\, x'_L + r_{12}\, y'_L + r_{13}\, f + r_{14}(f / z_L) &= x'_R\,(z_R / z_L) \\
r_{21}\, x'_L + r_{22}\, y'_L + r_{23}\, f + r_{24}(f / z_L) &= y'_R\,(z_R / z_L) \\
r_{31}\, x'_L + r_{32}\, y'_L + r_{33}\, f + r_{34}(f / z_L) &= f\,(z_R / z_L)
\end{aligned}
$$

And satisfies the 6 orthonormality constraints

Scale Ambiguity

Same image coordinates can be generated by doubling $\overline{r_L}\ \overline{r_R}\ \overline{r_0}$



Hence, we can find $\overline{r_0}$ only upto a scale factor!

So, fix scale by using constraint: $\overline{r_0} \cdot \overline{r_0} = 1$ (1 additional equation)

How many scene points are needed?

If we have n pairs of image coordinates:

Number of equations: $3n + 6 + 1 = 3n + 7$

Number of unknowns: $2n + 12$

$n = 5$ points will give us equal number of equations & unknowns.

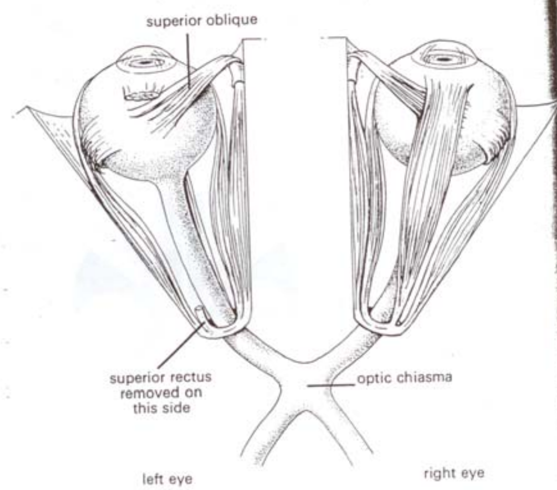
In theory, 5 points indeed are good enough if chosen carefully.

However, in practice, more points are used and
an over-determined system of equations is solved as before.

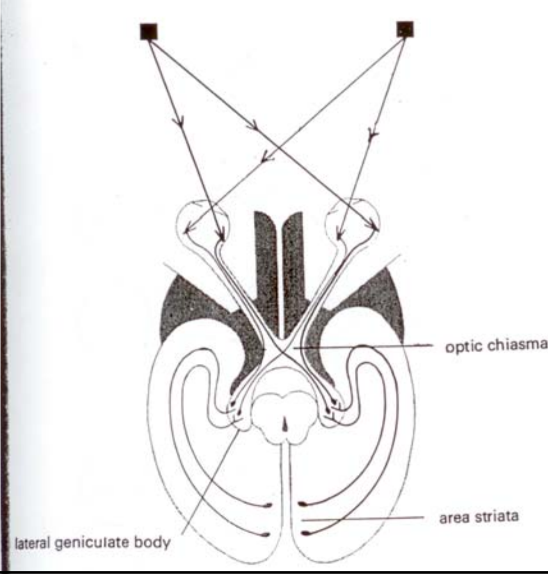
Eye and the Brain

FROM "EYE & BRAIN"
GREGORY

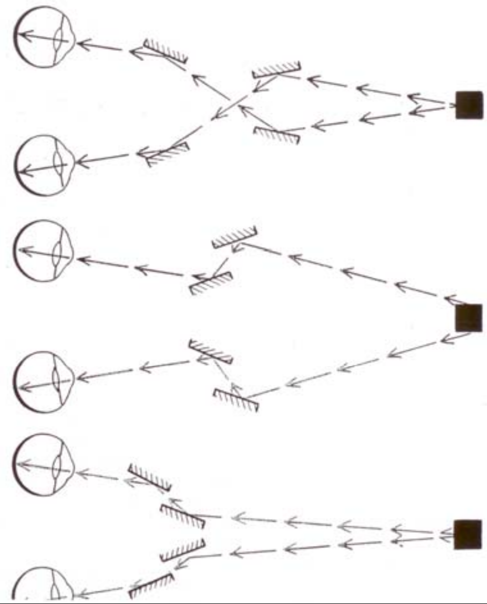
5.7 The muscles which move the eye. The eyeball is maintained in position in the orbit by six muscles, which move it to direct the gaze to any position and give convergence of the two eyes for depth perception. They are under continuous tension and form a delicately balanced system, which when upset can give illusions of movement.



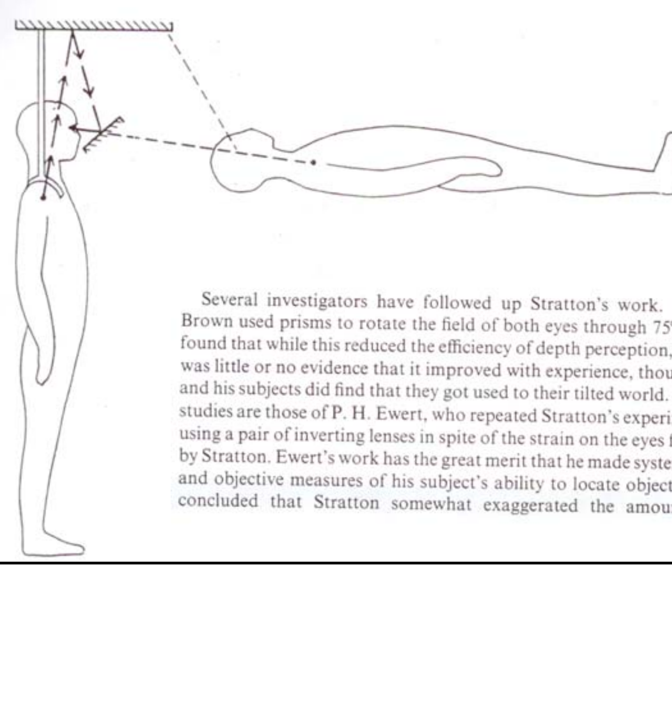
4.5 The optic pathways of the brain. The optic nerve divides at the chiasma, the right half of each retina being represented on the right side of the occipital cortex, the left side on the left half. The lateral geniculate bodies are relay stations between the eyes and the visual cortex.



5.12 Switching the eyes with mirrors. (Top) A *pseudoscope* – gives reversed depth, but only when depth is somewhat ambiguous. (Centre) A *telestereoscope* – effectively increases the separation of the eyes. (Bottom) An *iconoscope* – reduces the effective distance apart of the eyes. These arrangements are all useful for studying convergence and disparity in depth perception.



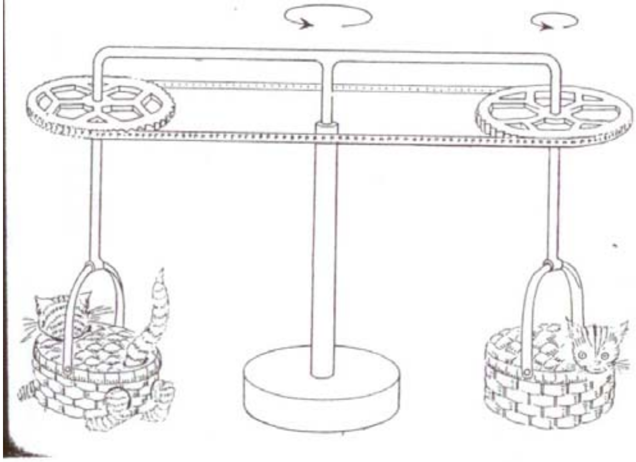
11.6 Stratton's experiment, in which he saw himself suspended in space before his eyes, in a mirror. He went for country walks wearing this arrangement.



Several investigators have followed up Stratton's work. G. C. Brown used prisms to rotate the field of both eyes through 75°, and found that while this reduced the efficiency of depth perception, there was little or no evidence that it improved with experience, though he and his subjects did find that they got used to their tilted world. Later studies are those of P. H. Ewert, who repeated Stratton's experiment, using a pair of inverting lenses in spite of the strain on the eyes found by Stratton. Ewert's work has the great merit that he made systematic and objective measures of his subject's ability to locate objects. He concluded that Stratton somewhat exaggerated the amount of

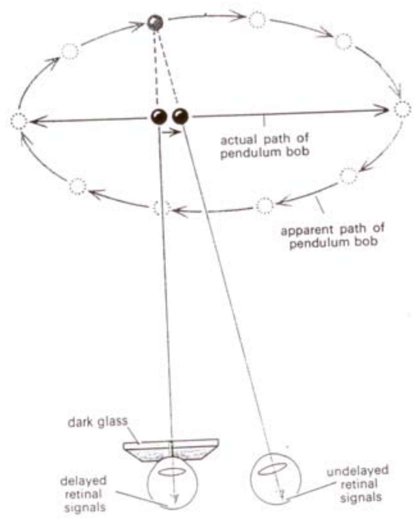
11.8 Pfister's hen wearing prisms to deviate light entering its eyes.

11.9 Apparatus designed by Held and Hein to discover whether perceptual learning takes place in a passive animal. The kitten on the right is carried about by the active kitten on the left. They thus have similar visual stimulation. Following visual experience limited to this situation, only the active animal is able to perform visual tasks – the passive animal remains effectively blind.



STEREOSCOPIC ILLUSION

6.3 The Pulfrich Pendulum. A pendulum swinging in a straight arc across the line of sight is viewed with a dark glass over one eye, both eyes being open. It appears to swing in an ellipse. This is due to the signals from the eye which is dark-adapted by the filter being delayed. The bob's increasing velocity towards the centre of its swing gives increasing *signalled* disparity: accepted as stereo depth signals corresponding to an ellipse.



Next Class

- Optical Flow and Motion
- Reading: Horn, Chapter 12.