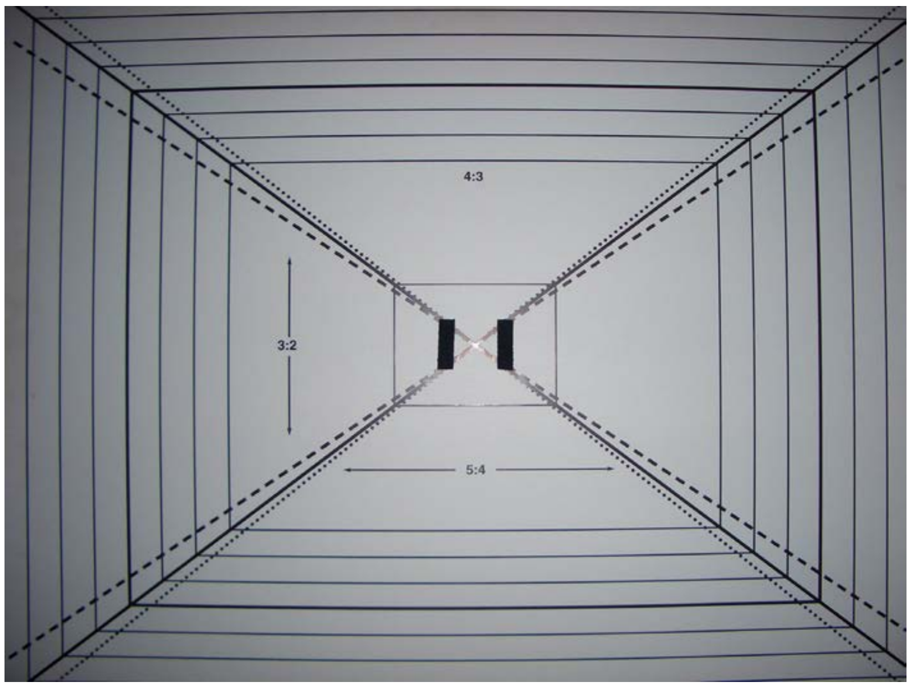
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ELECTRICAL ENGINEERING

6.801/6.866 MACHINE VISION



Problem 1: Quick way of estimating (the first coefficient of) radial distortion. A simple model of radial distortion is provided by the equation

$$r' = r(1 + k_2 r^2 + \ldots)$$

where r is the undistorted distance of a point in the image from the center, while r' is the distorted distance. We will ignore higher order coefficients and estimate the first coefficient k_2 of radial distortion from an image of a square lying in a plane perpendicular to the optical axis. Note that negative k_2 leads to "drum barrel" distortion, while positive k leads to "pin cushion" distortion (Wide angle lenses typically exhibit "drum barrel" distortion).

Suppose that the image of the square is centered in the image plane and that in the absence of distortion, the image of a corner of the square would be a distance r_c from the center, while the middle of one of the sides of the image of the square would be a distance r_m from the center. Next, assume that in the presence of radial distortion these distances are r_c' and r_m' respectively. Note that we directly observe only r_c' and r_m' — we cannot directly measure r_c or r_m.

Clearly $r_c = \sqrt{2} r_m$ (and this is also what we would observe in the absence of radial distortion).

(a) The difference $r_c' - \sqrt{2} r_m'$ is a quantity that varies with distortion and can be used to estimate k_2. Show that
$$r_c' - \sqrt{2} r_m' = \sqrt{2} k_2 r_m^3$$

(b) We could use this formula to estimate k_2 if only we knew r_m. Show that
$$r_m = 2r_m' - \frac{1}{\sqrt{2}} r_c'$$
How then can you determine the coefficient of radial distortion k_2?

(c) Suppose measurements yield $r_m' = 90$ and $r_c' = 113.137$. Calculate the difference in part (a) above. Then estimate k_2, as well as r_m and r_c.

Problem 2: Here we add scale to absolute orientation (but only in 2D, for simplicity). Consider a flying robotic system that uses binocular stereo to obtain three-dimensional information from pairs of images. Suppose that the scale of the recovered three-dimensional coordinates is not known accurately because the baseline between exposure stations is not known with precision. Now suppose that two such three-dimensional models — obtained along different flight paths — are to be related. In this case, determining the absolute orientation requires that, in addition to translation and rotation, a scale factor relating the two three-dimensional models be found as well.

Here we explore — in a two-dimensional version of the problem — the implications of chosing different ways of introducing a scale factor into the error expression to be minimized. Consider 'left' coordinates $\mathbf{l}_i$:
$$(a, 0), \quad (0, b), \quad (-a, 0), \quad (0, -b)$$
and the corresponding 'right' coordinates $\mathbf{r}_i$:
$$(c, 0), \quad (0, d), \quad (-c, 0), \quad (0, -d)$$
of four points (where $\mathbf{l}_1 = (a, 0)$ corresponds to $\mathbf{r}_1 = (c, 0)$ etc.). Hint: It may help to draw two diagrams showing the two sets of measurements.

(a) What is the best fit translation from 'left' to 'right' coordinate system? (Hint: this can be determined without any detailed calculation)

(b) What is the best fit rotation from 'left' to 'right' coordinate system? (Hint: this can be determined without any detailed calculation)

(c) Suppose that we express the coordinate transformation in the form
$$\mathbf{r}_i = s_1 R(\mathbf{l}_i) + \mathbf{t}$$
where $R(\ldots)$ is the rotation, $\mathbf{t}$ is the translation and s_1 is a scale factor. Show that minimizing the sum of squares of the magnitudes of the errors

$\mathbf{r}_i - \big(s_1 R(\mathbf{l}_i) + \mathbf{t}\big)$ leads to

$$s_1 = \frac{ac + bd}{a^2 + b^2}$$

(Hint: use the known rotation and translation to simplify the expression).

(d) Suppose that we express the coordinate transformation instead in the form

$$\mathbf{l}_i = s_2 R(\mathbf{r}_i) + \mathbf{t}$$

where s_2 is a scale factor. Show that minimizing the sum of squares of the magnitudes of the errors $\mathbf{l}_i - \big(s_2 R(\mathbf{r}_i) + \mathbf{t}\big)$ leads to

$$s_2 = \frac{ca + db}{c^2 + d^2}$$

(Hint: use the rotation and translation determined above to simplify the expression). When is $s_2 = 1/s_1$? What is the sum of squares of errors then?

(e) The two results above illustrate asymmetries between the way the 'left' and 'right' coordinates are treated, and the way they appear in the resulting expression for the scale factor. Now consider instead

$$\frac{1}{\sqrt{s_3}}\mathbf{r}_i = \sqrt{s_3}\,R(\mathbf{l}_i) + \mathbf{t}$$

Find s_3 that minimizes the sum of $\big\|(1/\sqrt{s_3})\mathbf{r}_i - \big(\sqrt{s_3}\,R(\mathbf{l}_i) + \mathbf{t}\big)\big\|^2$. Show that s_3 is the square root of s_1/s_2. Do correspondences between coordinates measured in the two coordinate systems need to be known in order to recover the scale factor s_3?

Problem 3: When comparing an object against a library of objects, or when determining the attitude of an object in space, it is useful to have an even sampling of the space of rotations. The rotation groups of the Platonic solids provide convenient uniform sampling of the space of rotations.

Consider the rotations of a tetrahedron that brings faces, edges and vertices into alignment. We'd like to express these in terms of unit quaternions. It helps to line up the tetrahedron with the coordinate axes in a symmetric way:

(a) Suppose the four vertices are at $(a, 0, b)$, $(a, 0, -b)$, $(-a, b, 0)$, and $(-a, -b, 0)$. For what values of a and b are these four vectors from the centroid of the tetrahedron (i) unit vectors, and (ii) at equal angles from one another?

(b) Some of the rotations of interest are those about lines from the centroid to the vertices (i.e. the four vectors in part(a)). These rotations are through angles of $\pm 2\pi/3$. Give the components of two quaternions that correspond to rotation through $2\pi/3$ about two *different* vectors.

(c) Now take the products of these two quaternions — in both possible orders — to generate two more rotations (remember that multiplication here does not commute).

(d) Finally, take the transitive closure. That is keep on multiplying the rotations you have generated pairwise with one another until no new rotations are generated. How many different rotations are there all together? How many of these are through angles of $\pm 2\pi/3$? Are there any rotations through angles other than 0 and $\pm 2\pi/3$ (Hint: you can reduce the amount of work by remembering that $-\mathring{q}$ represents the same rotation as $\mathring{q}$).

Problem 4: Relative Orientation. Imagine that we didn't know about the virtues of representing rotations using quaternions and we wished to solve the relative orientation problem using orthonormal matrix notation instead.

(a) Show that the coplanarity condition $[\mathbf{r}\,\mathbf{b}\,\mathbf{l}'] = 0$ (where $\mathbf{l}'$ is the 'left' vector rotated into the 'right' coordinate system) can be written

$$\mathbf{r}^T\mathbf{E}\mathbf{l} = 0$$

where $\mathbf{E} = \mathbf{B}\mathbf{R}$ is the so-called "essential matrix," with $\mathbf{R}$ a 3×3 orthonormal rotation matrix, and $\mathbf{B}$ the skew symmetric matrix corresponding to taking the cross-product with the baseline. (That is, $\mathbf{B}\mathbf{v} = \mathbf{b}\times\mathbf{v}$ for any $\mathbf{v}$.) Show that $\det(\mathbf{E}) = 0$ (Hint: find a non-zero vector $\mathbf{v}$ such that $\mathbf{E}\mathbf{v} = 0$).

Next, let's assume that the matrix $\mathbf{E}$ has already been estimated from correspondences between the left and right images and focus on recovering the baseline and the rotation from the "essential matrix." In essence, we have to split $\mathbf{E}$ into the product of a skew-symmetric matrix ($\mathbf{B}$) and an orthonormal matrix ($\mathbf{R}$).

(b) Show that each column of the "essential matrix" is orthogonal to the baseline. How can you easily obtain the direction of the baseline from any two columns of the matrix?

(c) Is the length of the baseline fixed by the "essential matrix"? Show that, if a set of corresponding ray directions satisfies the "essential matrix" constraint $\mathbf{r}^T\mathbf{E}\mathbf{l} = 0$, then they also satisfy $\mathbf{r}^T\mathbf{E}'\mathbf{l} = 0$ where $\mathbf{E}' = k\mathbf{E}$.

(d) It is often convenient to separate the recovery of translation from that of rotation. Show that $\mathbf{E}\mathbf{E}^T = -\mathbf{B}^2$ (and hence independent of rotation). Then show that $\mathbf{B}^2 = \mathbf{b}\mathbf{b}^T - (\mathbf{b}\cdot\mathbf{b})\mathbf{I}$, and that $\mathrm{Trace}(\mathbf{B}^2) = -2(\mathbf{b}.\mathbf{b})$. Conclude that

$$\mathbf{b}\mathbf{b}^T = (1/2)\mathrm{Trace}(\mathbf{E}\mathbf{E}^T)\mathbf{I} - \mathbf{E}\mathbf{E}^T$$

where $\mathrm{Trace}(\mathbf{E}\mathbf{E}^T)$ is just the sum of squares of the elements of $\mathbf{E}$. This provides another way to recover the baseline. How can one get $\mathbf{b}$ from $\mathbf{b}\mathbf{b}^T$? Are the magnitude and sign of $\mathbf{b}$ uniquely determined?

Problem 5: When we view the corner of a rectangular building, we obtain three edges in the image from which we can determernmine the viewing direction in the coordinate system defined by the three sets of parallel edges in the building. To simplify matters, we'll consider orthographic projection (or equivalently, that the camera is aimed so that the image of the corner falls in the center of the image).

So, consider an *orthographic* projection of the corner of a rectangular building. The task is to recover the orientation of the rectangular "brick" with respect to the image plane from the angles measured in the image between the lines meeting at the vertex (for convenience, place the origin at the vertex). Let the ends (arbitrarily defined) of the three lines be given by the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ — which of course all lie in the image plane. (The z component of these three vectors is zero since the z-axis is taken to be the direction of projection and hence perpendicular to the image plane.)

The image line $\mathbf{a}$ corresponds to an edge of the brick parallel to $\mathbf{a}' = \mathbf{a} + a\,\hat{\mathbf{z}}$, for some unknown a (since in orthographic projection $\mathbf{a}$ is obtained from $\mathbf{a}'$ by dropping the z component). Similarly the image line $\mathbf{b}$ corresponds to an edge parallel to $\mathbf{b}' = \mathbf{b} + b\,\hat{\mathbf{z}}$, while the image line $\mathbf{c}$ corresponds to an edge parallel to $\mathbf{c}' = \mathbf{c} + c\,\hat{\mathbf{z}}$.

(a) Find the unknown scalars a, b, and c. Hint: Use the fact that the edges of the brick are supposed to be orthogonal, and hence so are $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$.

(b) Let α be the angle that the vector $\mathbf{a}'$ makes with the image plane (and hence with $\mathbf{a}$). Similarly, let β be the angle between $\mathbf{b}'$ and the image plane and γ be the angle between $\mathbf{c}'$ and the image plane. Show that

$$\tan\alpha = \sqrt{-\frac{\cos B \cos C}{\cos A}}$$

where A is the angle in the image plane between the lines $\mathbf{b}$ and $\mathbf{c}$, B is the angle between the lines $\mathbf{c}$ and $\mathbf{a}$, while C is the angle between $\mathbf{c}$ and $\mathbf{a}$. Give similar expressions for $\tan\beta$ and $\tan\gamma$.

(c) What constraints — if any — on the angles A, B, and C are imposed by the fact that the terms under the square-root sign must be non-negative?