Massachusetts Institute of Technology

Department of Computer Science and Electrical Engineering

6.801/6.866 Machine Vision

Problem 1: Rotational Optical Flow. In our analysis of the "fixed flow" problem we have so far considered only translational motion. We can generalize this to the case where there is in additon in-plane rotation. The whole image translates with unkown velocity (u, v) and with an in-plane *rotational* component ω. That is, in addition to translating with velocity (u_0, v_0), say, the image is also rotating with angular velocity ω (rad/sec) about some (known) reference point (x_0, y_0) (which could be chosen to be the principal point, or just the center of the image, but here we keep it general).

(a) Show that the velocity (u, v) at point (x, y) in the image is

$$u = u_0 - \omega y \quad \text{and} \quad v = v_0 + \omega x$$

where, for convenience, we have defined $x = (x - x_0)$ and $y = (y - y_0)$.

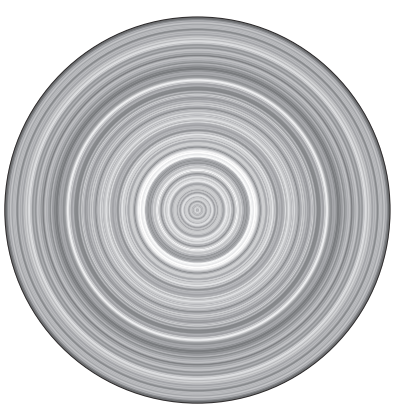
(b) Insert u and v (now spatially varying rather than fixed for the whole image) into the brightness change constraint equation.

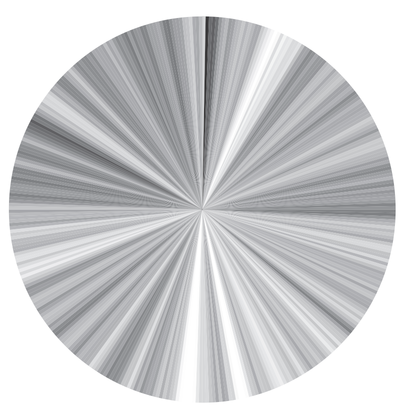
(c) The task is to find the values of the unknown parameters u_0, v_0, and ω that minimize the integral of the errors squared:

$$\iint (uE_x + vE_y + E_t)^2 \, dx \, dy$$

You should obtain three linear equations in the three unkonwns.

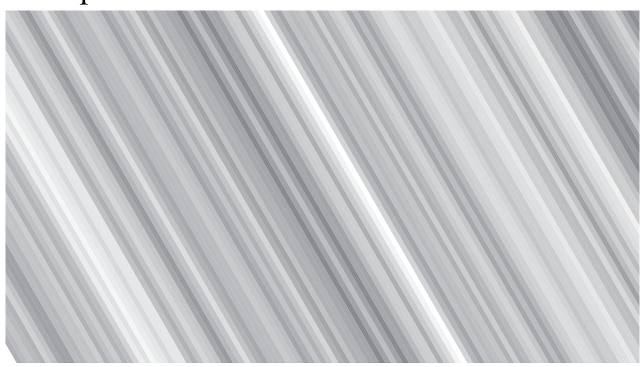
(d) In order to obtain a unique solution, what conditions must the determinant of the 3×3 coefficient matrix satisfy?





(e) What happens when the brightness pattern is radially symmetric about the point (x_0, y_0) (above left)? Show that in this case $(x\ , y\ ).(E_y, -E_x) = 0$. Can one still recover the translational component of motion (u_0, v_0)? Explain.

(f) What happens when the brighntess pattern is constant along radii, much like a pie chart (above right)? Do you expect to be able to recover translation and rotation? Explain.



(g) What happens when the brightness gradient everywhere has the same direction (that is $E_y = kE_x$) as in the image above? Can one still recover the rotational component of motion ω? Explain.

Problem 2: Time-to-Contact w.r.t. inclined surface. In class we analyzed a direct method for recovering "time to contact" (TTC) in translational motion towards a planar surface. In that example, the translational motion $(U, V, W)^T$ could be along any direction, but the plane had to be perpendicular to the optical axis, that is, $Z = Z_0$

Here we instead allow the plane to have *any* orientation, but restrain the motion to be along the optical axis, that is, $U = V = 0$. The equation of a plane is linear in X, Y and Z (coordinates in the camera-centric coordinate system). It can, for example, be written in the form

$$Z = Z_0 + pX + qY$$

where $p = \partial Z/\partial X$ and $q = \partial Z/\partial Y$ are the slopes of the surface in the X and Y directions, while Z_0 is the distance from the center of projection to where the optical axis pierces the plane.

With translational motion along the optical axis, we have simple radial optical flow with focus of expansion (FOE) at $(0, 0)$:

$$u = -(W/Z)x \quad \text{and} \quad v = -(W/Z)y$$

where $W = dZ/dt$ is the velocity of the planar surface in the Z direction (equivalently the negative of the motion of the camera relative to the planar surface).

(a) For our purposes, we need Z as a function of image coordinates x and y (rather than as a function of X and Y). Show that

$$Z\big(1 - p(x/f) - q(y/f)\big) = Z_0$$

(b) Next, using the brightness change constraint equation, show that

$$-(W/Z_0)\big(1 - p(x/f) - q(y/f)\big)G + E_t = 0$$

where $G = (x, y) \cdot (E_x, E_y)$

(c) Formulate a least squares problem for finding three unknown parameters

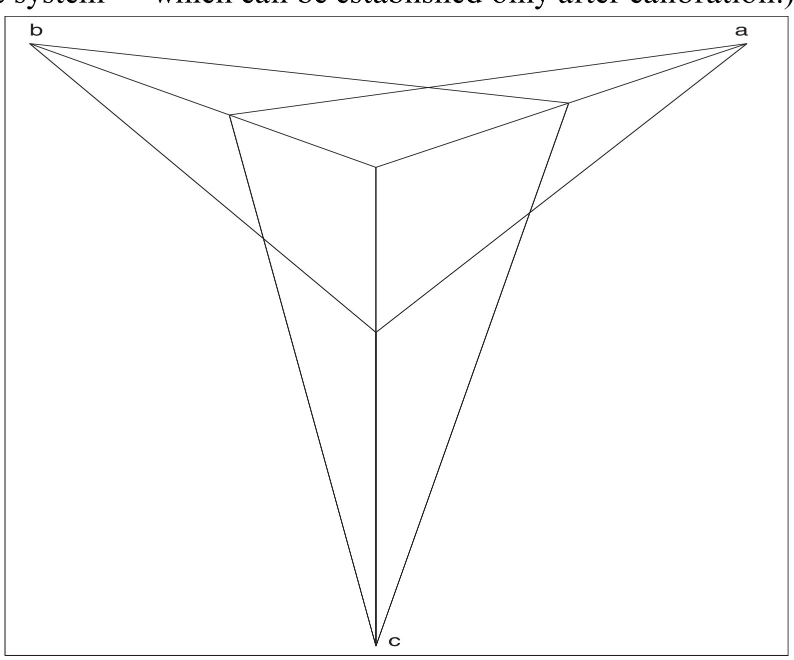
$$P = \frac{p}{f}\frac{W}{Z_0}, \quad Q = \frac{q}{f}\frac{W}{Z_0}, \quad \text{and} \quad C = -\frac{W}{Z_0}$$

that best fit the image sequence. Show that the minimum occurs for values of P, Q, C that satisfy three linear equations.

(d) How can you recover the surface slopes, p and q, from P, Q, and C? What is the time to contact, T, in terms of P, Q, and C?

Problem 3: Camera calibration using vanishing points. We can use a brick-shaped object for camera calibration. We do not need to know the size of the object, its orientation, or its distance from the camera. The twelve edges of a brick-shaped object come in three groups of parallel lines. The lines in each group are perpendicular to the lines in the other two groups. Each group of parallel lines creates a vanishing point in the image.

Suppose that we establish an image plane coordinate system based on the photodetector array column number (for x), the row number (for y) — and $z = 0$ for points in the image plane. (Note that that is *not* our standard "camera-centric" coordinate system — which can be established only after calibration.)



Suppose that the measured positions of the three vanishing points are $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ in that coordinate system Note that $\mathbf{a} \cdot \hat{\mathbf{z}} = 0$, $\mathbf{b} \cdot \hat{\mathbf{z}} = 0$ and $\mathbf{c} \cdot \hat{\mathbf{z}} = 0$ since these points lie in the image plane. Next, suppose that the center of projection is at $\mathbf{p} = (x_p, y_p, f)^T$ in the same coordinate system. We will recover the unknown position of the center of projection $\mathbf{p}$ from the positions $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ of the three vanishing points in the image plane.

(a) Given the figure above, explain why there are three sets of lines parallel to $(\mathbf{a} - \mathbf{p})$, $(\mathbf{b} - \mathbf{p})$, and $(\mathbf{c} - \mathbf{p})$ respectively.

(b) Explain why the following holds

$$(\mathbf{a} - \mathbf{p}) \cdot (\mathbf{b} - \mathbf{p}) = 0, \quad (\mathbf{b} - \mathbf{p}) \cdot (\mathbf{c} - \mathbf{p}) = 0, \quad (\mathbf{c} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{p}) = 0$$

(c) By subtracting equations pairwise, show that

$$(\mathbf{a} - \mathbf{p}) \cdot (\mathbf{c} - \mathbf{b}) = 0, \quad (\mathbf{b} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{c}) = 0, \quad (\mathbf{c} - \mathbf{p}) \cdot (\mathbf{b} - \mathbf{a}) = 0$$

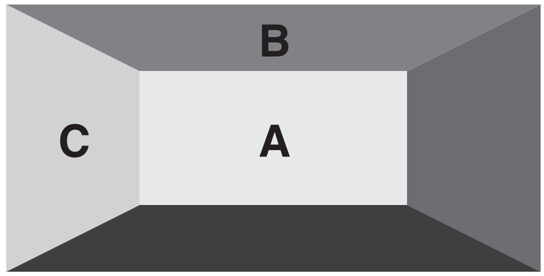
(d) Show that each of these three linear equations in $\mathbf{p}$ correspond to a plane with surface normal parallel to the image plane. Are the three equations always linearly independent? If not, when are they linearly dependent?

(e) How are these planes oriented relative to the image plane? How is the line of intersection of any pair of these planes related to the image plane?

(f) Suppose now that the measured positions of the three vanishing points in an image (Note: above figure not to scale) are $\mathbf{a} = (500, 500, 0)^T$, $\mathbf{b} = (100, 500, 0)^T$, and $\mathbf{c} = (300, 100, 0)^T$. By solving two linear equations for two unknowns, find the point on the line defined in part (e) that lies in the image plane. How is this point related to the principal point?

(g) We now know x_p and y_p, the first two components of $\mathbf{p}$, and so only need f to complete the calibration of the camera. Find the height f of $\mathbf{p}$ above the image plane (you may need to solve a quadratic equation). Conclude that the principal distance f is approximately $173.2\ldots$ pixels (expressed in units of spacing between pixels).

Problem 4: Source from shading. In photometric stereo we try to find surface orientaton, typically with knowledge of the light source position and surface reflectance properties. In this problem we try to "invert" this to find the light source position given surface orientations. In an image of a rectangular 'button' rising above a flat background, the brightness of the indicated regions is as follows: A has grey value 212, B has grey value 175, and C has grey value 200. Importantly, we are told that the bevelled edges of the button are inclined 45° with respect to the plane of the background (which is parallel to region A).



(a) Consider a coordinate system in the plane of the background, with x running to the right and y upwards. Let the z axis be perpendicular to the background plane, pointing "out of the page." Write the normals of each of the regions A, B, and C as vectors with three components.

(b) Next, assume that the surface has reflecting properties that closely match Lambert's 'law', that is, that brightness is proportional to the cosine of the incident angle. Using the three measurements of brightness, find the unit vector in the direction of the light source. (You may find yourself solving simultaneous equations, or, using guesswork and iterative refinement).

Problem 5: Shape from Shading.

(a) Consider the quadratic surface

$$z(x, y) = ax^2 + bxy + cy^2$$

Find an expression for the unit surface normal

$$\hat{\mathbf{n}}(x, y)$$

as a function of x and y. Assuming a single distance light source in direction

$$\hat{\mathbf{s}} = (s_x, s_y, s_z)^T$$

and Lambertian surface characteristics, find the surface radiance $L(x, y)$ as a function of x and y. The irradiance of the source on a plane perpendicular to the incident rays is E_0 ($\mathrm{W} \cdot \mathrm{m}^{-2}$).

(b) Consider a surface made of a material that has reflectance map

$$R(p, q) = p^2 + q^2$$

and results in an image

$$E(x, y) = 2 - \cos(2x) - \cos(2y)$$

show that the surface may have the shape

$$z = a\cos(x) + b\cos(y)$$

for suitable values of a and b. Are there other surfaces that give rise to the same image?