MASSACHUSETTS INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ELECTRICAL ENGINEERING

6.801/6.866 MACHINE VISION HWP 1 SOLUTIONS

Handed out: 2004 Sep. 16th

PROBLEM 1

Part (a,b).
If someone solved for u (using the least squares method) assuming that $uE_x + E_t = 0$, their answer would be

$$u^* = \frac{-\int E_x E_t dx}{\int E_x^2 dx} \tag{1}$$

as was shown in class. Now, since in our case we have an object whose brightness is changing with a constant speed k, we have that $uE_x + (E_t - k) = 0$, and so the right derivation of u would be like the one done in class, but it would use $E_t' = E_t - k$ instead of E_t, which gives:

$$u = \frac{-\int E_x E_t' dx}{\int E_x^2 dx} = \frac{-\int E_x (E_t - k) dx}{\int E_x^2 dx} = \frac{-\int E_x E_t dx}{\int E_x^2 dx} + \frac{\int k E_x dx}{\int E_x^2 dx} \tag{2}$$

Therefore, the erratic computation, i.e. u^*, the computation that did not take the constant change of brightness k into the account, fails to give us the proper u in general. Instead, it gives us

$$u^* = u - \frac{\int k E_x dx}{\int E_x^2 dx}$$

Now, notice that the error term is

$$\Delta = u^* - u = -\frac{k \int_{x_0}^{x_1} E_x dx}{\int E_x^2 dx} = -\frac{k(E(x_1, t) - E(x_0, t))}{\int E_x^2 dx}$$

Therefore, if the brightness at the right end (x_1) is the same as the brightness at the left end (x_0) of the image, Δ is zero, and hence $u^* = u$ and the "corrupted" computation introduces no error.

Some remarks about this problem:

1. In this problem we assumed a parallel projection, i.e. that the object at point in the world x is projected in the camera to point x on the image plane. This is a simplifying assumption which we'll rarely see again in this class.

2. The least square solution for u we got in class (and discussed above) could also be used over dt direction, i.e. instead of finding u that minimizes

$$\int_{x_1}^{x_2} (uE_x + E_t)^2 dx$$

at some point t, we could find u that minimizes

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} (uE_x + E_t)^2 dx dt$$

across some duration of our measurement from time t_1 to t_2. If we did that, all the calculations would be the same, except there will always be additional $\int_{t_1}^{t_2} (...) dt$ integration on top.

3. The special case considered in part **(b)**, i.e. that $E(x_1, t) = E(x_0, t)$, is not arbitrary: If you film an object that moves between point x_1 and x_2, and that object is completely contained within the field of vision of your camera, then at the edges of your view, at points x_1 and x_2 of the image, you will see the background (for most t). If that background has uniform brightness then indeed $E(x_1, t) = E(x_0, t)$.

Part (c).
Now we know that u and k are bound by the

$$uE_x + E_t - k = 0$$

constraint. From measurements of $E(x, t)$ over some range (x_1, x_2) and (t_1, t_2) (by neccesity our measurements will be discrete, but we can use them to approximate the continuous data), we can solve for u, k (unlike parts **(a,b)**, in this part k is unknown) using the least square method. In other words, we are searching for (u, k) that minimize

$$s(u, k) = \int_{t_1}^{t_2} \int_{x_1}^{x_2} (uE_x + E_t - k)^2 dx dt$$

(From now on, we'll drop the $\int_{t_1}^{t_2} (...) dt$ terms, see point **(2)** above.) Since it is a quadratic expression, it achieves a minimum when both $\frac{\partial s}{\partial u}$ and $\frac{\partial s}{\partial k}$ are zero. Taking these derivatives we get:

$$\begin{aligned} \frac{\partial s}{\partial u} &= \int_{x_1}^{x_2} (2uE_x^2 + 2E_x(E_t - k)) dx = 0 \\ \frac{\partial s}{\partial k} &= \int_{x_1}^{x_2} (2k - 2(uE_x + E_t)) dx = 0 \end{aligned}$$

which leads to the following two linear equations

$$\begin{array}{rcl} (\int E_x^2 dx)u \;+\; (-\int E_x dx)k &=& -\int E_x E_t dx \\ (-\int E_x dx)u \;+\; (\int dx)k &=& \int E_t dx \end{array}$$

I.e. we have $M * [u,k]^T = A^T$, and hence $[u,k]^T = M^{-1}A^T$. For square matrix M, computing M^{-1} is easy:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} * \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

and hence:

$$\begin{array}{rcl} u &=& \dfrac{-\int dx \int E_t E_x dx + \int E_x dx \int E_t dx}{\int dx \int E_x^2 dx - (\int E_x dx)^2} \\ k &=& \dfrac{-\int E_x dx \int E_t E_x dx + \int E_x^2 dx \int E_t dx}{\int dx \int E_x^2 dx - (\int E_x dx)^2} \end{array}$$

Part (d).
Since $\int_{x_1}^{x_2} E_x(x,t)dx = E(x_2,t) - E(x_1,t)$, then if we assume that $E(x_2,t) = E(x_1,t)$ then the $\int E_x dx$ term becomes zero everywhere, which greatly simplifies the above formulas for u and k to the following:

$$\begin{array}{rcl} u &=& \dfrac{-\int dx \int E_t E_x dx}{\int dx \int E_x^2 dx} = \dfrac{-\int E_t E_x dx}{\int E_x^2 dx} \\ k &=& \dfrac{\int E_x^2 dx \int E_t dx}{\int dx \int E_x^2 dx} = \dfrac{\int E_t dx}{\int dx} = \dfrac{\int E_t dx}{x_2 - x_1} \end{array}$$

It is good to see that in case of $E(x_2,t) = E(x_1,t)$, when we get u by solving for best-fit for both u and k, we get the same answer as we would if we knew k and just solved for the best fit for u, which was the result of parts **(a,b)** above.

PROBLEM 2

Part (a).
$b = a\cos\theta$ and $e = \sin\theta$.

Part (b).
Take $\frac{db}{d\theta} = -a\sin\theta = -ae$. Therefore the error sensitivity $\frac{d\theta}{db} = \frac{1}{-ae}$.

Part (c).
No, it is not a good position, because if you put a camera on the line passing through the axis of the wheel then θ will be about zero, which will make $e \approx 0$, and thus the error sensitivity $\frac{1}{-ae}$ will be sky-high. Instead, you should position the camera so that $\frac{1}{-ae}$ is as small as possible, i.e. when e is large, i.e. when $\theta \approx \pi/2$, i.e. the camera should be placed along the plane of the rotating wheel, not its axis.

PROBLEM 3

Part (a).
Let r_1, r_2 be the distances respectively from the left and the right camera to the image. Notice that:

$$x + b/2 = r_1 \sin\theta_1 \tag{3}$$

and

$$z = r_1 \cos\theta_1 \tag{4}$$

and that from the law of sines we have:

$$\frac{r_1}{\sin\theta_2} = \frac{r_2}{\sin\theta_1} = \frac{b}{\sin(\theta_1 - \theta_2)} \tag{5}$$

From eq (4) and (5) we immediately get

$$z = b\frac{\cos\theta_1 \cos\theta_2}{\sin(\theta_1 - \theta_2)}$$

And from eq (3) and (5) we have that

$$\begin{aligned} x &= -\frac{b}{2} + \frac{b\cos\theta_2 \sin\theta_1}{\sin(\theta_1 - \theta_2)} \\ &= \frac{b}{2}\,\frac{-\sin(\theta_1 - \theta_2) + 2\cos\theta_2\sin\theta_1}{\sin(\theta_1 - \theta_2)} \\ &= \frac{b}{2}\,\frac{\sin(\theta_2 + \theta_1)}{\sin(\theta_1 - \theta_2)} \end{aligned}$$

If $\theta_1 \approx 0$ then

$$\begin{aligned} x &= -\frac{b}{2} \\ z &= -b\frac{\cos\theta_2}{\sin\theta_2} \end{aligned}$$

If $\theta_2 \approx 0$ then

$$\begin{aligned} x &= \frac{b}{2} \\ z &= b\frac{\cos\theta_2}{\sin\theta_2} \end{aligned}$$

Part (b).
If $\theta_1 \approx 0$ and $\theta_2 \approx 0$ then $\cos\theta_1 \approx \cos\theta_2 \approx 1$, so from the equation above we get:

$$z \approx \frac{b}{\sin(\theta_2 - \theta_1)}$$

Part (c).
Taking the partial derivative of the above approximation for z with respect to θ_1 and θ_2, we get:

$$\begin{aligned} \frac{\partial z}{\partial \theta_1} &\approx \frac{-b\cos(\theta_1 - \theta_2)}{\sin^2(\theta_1 - \theta_2)} \\ \frac{\partial z}{\partial \theta_2} &\approx \frac{b\cos(\theta_1 - \theta_2)}{\sin^2(\theta_1 - \theta_2)} \end{aligned}$$

Since $\cos(\theta_2 - \theta_1) \approx \cos(0) \approx 1$ for $\theta_1 \approx 0$ and $\theta_2 \approx 0$, we get:

$$\frac{\partial z}{\partial \theta_1} \approx \frac{-b}{\sin^2(\theta_1 - \theta_2)} \approx -\frac{z^2}{b}$$

and similarily, $\frac{\partial z}{\partial \theta_2} \approx \frac{z^2}{b}$.

Therefore, z grows as $\frac{z^2}{b}$ relative to changes in θ_1 and θ_2. Consequently, the relative error dz/z grows as $\frac{z^2}{b}/z = \frac{z}{b}$.

PROBLEM 4

Set t to be the length of a side on the pyramid's base. The three edges leading from the apex of the pyramid to the base can then be described by the vectors $(\frac{t}{2}, \frac{-t\sqrt{3}}{6}, -h)^T$, $(\frac{-t}{2}, \frac{-t\sqrt{3}}{6}, -h)^T$, and $(0, \frac{t\sqrt{3}}{3}, -h)^T$. From this, we can take cross products to calculate the direction of three normal vectors for each of the facets:

$$
\begin{aligned}
\mathbf{A} &= (\frac{ht\sqrt{3}}{2}, \frac{ht}{2}, \frac{t^2}{2\sqrt{3}})^T \\
\mathbf{B} &= (0, -ht, \frac{t^2}{2\sqrt{3}})^T \\
\mathbf{C} &= (\frac{-ht\sqrt{3}}{2}, \frac{ht}{2}, \frac{t^2}{2\sqrt{3}})^T
\end{aligned}
$$

Note that $|A| = |B| = |C|$, thus, the unit normal vectors of the surface are:

$$
\begin{aligned}
\mathbf{n_A} &= k(\frac{ht\sqrt{3}}{2}, \frac{ht}{2}, \frac{t^2}{2\sqrt{3}})^T \\
\mathbf{n_B} &= k(0, -ht, \frac{t^2}{2\sqrt{3}})^T \\
\mathbf{n_C} &= k(\frac{-ht\sqrt{3}}{2}, \frac{ht}{2}, \frac{t^2}{2\sqrt{3}})^T
\end{aligned}
$$

where $k = 1/|A|$.

Based on the Lambert's law, the observed brightness is porportional to the dot products of the normal vectors of each side with the source vector $V = (1/\sqrt{2}, 0, 1/\sqrt{2})^T$. Since the three triangles are coated with same materials, the albedo is the same for three sides.

Therefore,

$$
\begin{aligned}
E_A &= \alpha A * V &= \alpha(\frac{ht\sqrt{3}}{2\sqrt{2}} + \frac{t^2}{2\sqrt{6}}) \\
E_B &= \alpha B * V &= \alpha(\frac{t^2}{2\sqrt{6}}) \\
E_C &= \alpha C * V &= \alpha(\frac{-ht\sqrt{3}}{2\sqrt{2}} + \frac{t^2}{2\sqrt{6}})
\end{aligned}
$$

where α is a constant.

Since E_A and E_C are in the ratio $3:1$, we can simply solve the equation:

$$
\begin{aligned}
\frac{-3ht\sqrt{3}}{2\sqrt{2}} + \frac{3t^2}{2\sqrt{6}} &= \frac{ht\sqrt{3}}{2\sqrt{2}} + \frac{t^2}{2\sqrt{6}} \\
\frac{2t^2}{2\sqrt{6}} &= \frac{4ht\sqrt{3}}{2\sqrt{2}} \\
\frac{t}{\sqrt{3}} &= 2h\sqrt{3} \\
\frac{t}{6} &= h
\end{aligned}
$$

And thus the height of the pyramid is 1/6 the length of a side.

ATTN: If you do not mention $|A| = |B| = |C|$ and same albedo condition, you will not receive full credit.

PROBLEM 5

Part (a).

The edges of the rectangular brick form three groups of parallel lines. The directions of these three groups of lines are orthogonal. Consequently, lines connecting the center of projection with the vanishing points (which are parallel to the corresponding edges of the rectangular brick) must also be orthogonal.

That is, $\mathbf{r} - \mathbf{a}$, $\mathbf{r} - \mathbf{b}$, and $\mathbf{r} - \mathbf{c}$ are orthogonal, and therefore

$$(\mathbf{r} - \mathbf{a}) \cdot (\mathbf{r} - \mathbf{b}) = 0$$

$$(\mathbf{r} - \mathbf{b}) \cdot (\mathbf{r} - \mathbf{c}) = 0$$

$$(\mathbf{r} - \mathbf{c}) \cdot (\mathbf{r} - \mathbf{a}) = 0$$

Part (b).

Expanding the dot products and subtracting equations pairwise we get

$$(\mathbf{r} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{c}) = 0$$

$$(\mathbf{r} - \mathbf{b}) \cdot (\mathbf{c} - \mathbf{a}) = 0$$

$$(\mathbf{r} - \mathbf{c}) \cdot (\mathbf{a} - \mathbf{b}) = 0$$

(This set of equations is redundant since the last equation can be obtained by subtracting the second from the first.)

Part (c).
Let $\mathbf{p}$ be the principal point (point where the perpendicular dropped from the center of projection strikes the image plane). Then by construction $\mathbf{p} - \mathbf{r}$ is perpendicular to the image plane, and hence perpendicular to any vector parallel to the image plane, including $\mathbf{b} - \mathbf{c}$, $\mathbf{c} - \mathbf{a}$ and $\mathbf{a} - \mathbf{b}$. We have $\mathbf{p} - \mathbf{a} = (\mathbf{p} - \mathbf{r}) - (\mathbf{r} - \mathbf{a})$, so

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

Similarly $(\mathbf{p} - \mathbf{b}) \cdot (\mathbf{c} - \mathbf{a}) = 0$ and $(\mathbf{p} - \mathbf{c}) \cdot (\mathbf{a} - \mathbf{b}) = 0$.

So a line connecting the principal point to one of the vertices of the triangle is orthogonal to the opposite edge of the triangle. We conclude that the principal point is at the 'orthocenter' of the triangle.

Part (d).
The principle distance is the distance of the center of projection from the image plane. The center of projection must lie on each of the three spheres that have the edges of the triangle as their diameters. Consequently it cannot be higher above the image plane than one half the shortest edge of the image plane triangle (i.e. the principal distance cannot be greater than the radius of any of the corresponding sphere).

Part (e).
Consider a cube cut by a plane passing through three vertices that are adjacent to a chosen vertex. Drop a perpendicular from the chosen vertex to this plane. This perpendicular corresponds to the optical axis and its length is the principle distance. The image triangle corresponds to the triangle cut out of the plane by the cube. Now consider the triangle formed by the points $(1,0,0)^T$, $(0,1,0)^T$, and $(0,0,1)^T$. Dropping a perpendicular onto this plane from the origin yields the point $(1,1,1)^T/3$, which lies at distance $1/\sqrt{3}$ from the origin. The length of each side of the equilateral triangle is $\sqrt{2}$. Consequently the ratio of principle distance to the edge length is $k = 1/\sqrt{6}$.

You can see that the upper bound of part **(d)** was not so bad. Here all edges have minimal length, and the bound of part **(d)** would limit f as no more than $1/2l$. The real $f = 1/\sqrt{6}l \approx 0.41l$, so the bound of part **(d)** is only 22% more than the real value. Furthermore, since you can probably argue that in the case of an equilateral triangle $k = f/l_{min}$ (i.e. the proportion of f to the minimal side of the triangle) is smallest, it would seem that the real f lies always between $0.41\,l_{min}$ and $0.5\,l_{min}$.