CS 498/698: Assignment 1, Image Formation and Lighting
Due: 11:30am, Tues. Jan. 31, 2006.

1. Perspective projection [4 marks]

(a) Trucco and Verri, Question 2.3. [2 marks]

Show that, in the pinhole camera model, three collinear points in 3-D space are imaged onto three collinear points on the image plane.

Recall the equations of perspective projection:

$$x = \frac{f}{Z}X; \qquad y = \frac{f}{Z}Y$$

where $\mathbf{P} = (X, Y, Z)^T$ is the 3-D point and $\mathbf{p} = (x, y)^T$ is the 2-D point, and f is the focal length.

(b) Horn, Exercise 2-12. [2 marks]

Straight lines in the three-dimensional world are projected as straight lines into the two-dimensional image. The projections of parallel lines intersect at a *vanishing point* in the image plane. When does the vanishing point of a family of parallel lines lie at infinity?

In the case of a rectangular object, a great deal of information can be recovered from lines in the images and their intersections. The edges of a rectangular solid fall into three sets of parallel lines, and so give rise to three vanishing points. In technical drawing one speaks of one-point, two-point, and three-point perspective. These terms apply to the cases in which two, one, or none of the three vanishing points lie at infinity. What alignment between the edges of the object and the image plane applies in each case? (Draw a diagram and/or describe in words.)

2. Combining images of different exposure [8 marks]

Finding the camera response function. Suppose we have P images taken at different exposure times $\Delta t_j, j = 1 \ldots P$. Let I_{ij} be the brightness of pixel i in image j. In each image

$i = 1 \ldots N$ pixels are measured. The brightness of each pixel is

$$I_{ij} = f(E_i \Delta t_j)$$

where I is the image brightness, E is the (unknown) scene brightness, and $f()$ is the (unknown) camera response function that maps the image brightness onto a discrete set of pixel values. Note that f is a function of the product of the scene brightness and the exposure time. That is, if the exposure time is doubled, the input to f doubles. The image brightness, however, is generally a nonlinear function of scene brightness.

Given a series of measurements, we solve for f as follows:

$$f^{-1}(I_{ij}) = E_i \Delta t_j$$

$$\ln f^{-1}(I_{ij}) = \ln E_i + \ln \Delta t_j$$

Letting $g = \ln f^{-1}$, gives

$$g(I_{ij}) = \ln E_i + \ln \Delta t_j$$

where $i = 1 \ldots N$ (pixels) and $j = 1 \ldots P$ (images).

To solve for g we select some subset of pixels (eg., 200 points) in all frames and minimize the following quadratic function.

$$\varepsilon^2 = \sum_{i=1\ldots N} \sum_{j=1..P} \left[w(I_{ij}) \left(g(I_{ij}) - \ln E_i - \ln \Delta t_j\right)\right]^2 + \lambda \sum_{z=2}^{n-1} \left[w(z) g''(z)\right]^2$$

where $g''(z) = g(z-1) - 2g(z) + g(z+1)$ and n is the number of grayscale levels in the image. The first term fits the camera response function, while the second term acts as a smoothing term that prefers functions with low curvature. $\lambda > 0$ trades off the effects of these two terms. A value of $\lambda = 100$ was used in our experiments. w is a weighting function that emphasizes the fit when pixels are near the midpoint of their range.

$$w(z) = \begin{cases} n - z & \text{if } z > n/2 \\ z & \text{if } z \leq n/2 \end{cases}$$

Finally, we add a constraint $g(n/2) = 0$. This constrains the function g to be zero at the midpoint of its domain. This is necessary since the above equations only determine g up to a constant factor.

The above function is minimized using the singular value decomposition. The result is a camera response function $g(z)$, $z = 1 \ldots n$ and a set of brightness values $\ln E_i$, $i = 1 \ldots N$.

(a) Equation counting [2 marks]

Write the number of *equations* and *unknowns* for the above equations in terms of N, P, and n. Compute these numbers when $N = 200$, $P = 3$, and $n = 256$, the case in our experiments below.

Combining images of different exposure. Since we can only determine the image response function g up to a scale factor, we let $\Delta t_j = 1$ for a reference image.

For each image I_j, we can compute the estimated scene brightness:

$$\hat{E}_j = \frac{1}{\Delta t_j} f^{-1}(I_j) = \frac{1}{\Delta t_j} e^{g(I_j)}$$

Using all images, we estimate the brightness:

$$\hat{E} = \sum_{j=1\ldots P} \frac{w(I_j)\hat{E}_j}{Z}$$

where $Z = \sum_j w_j(I_j)$ and w is the pixel weighting function: (Note: w, E, I_j, and Z are all functions of pixel position (x, y).)

To display the composite image, we plot $\log \hat{E}$. Note that since this accentuates the low intensity information, it is not directly comparable with the input images I_j.

(b) Experiments [6 marks]

In the program, `wyckoff.m`, I use three images, taken with *exposure values* (EV) as follows: (2, 1, $\frac{1}{2\sqrt{2}}$). A unit change in exposure value corresponds to a doubling or halving of exposure times. It may be achieved by either changing the exposure time, the camera aperature, or both. Please do the following experiments:

1. Run the program, print out, and comment on the composite image. In particular, what details does it capture that are not visible in the input images?
2. Modify the program to compute $\hat{E}_j$, $i = 1\ldots3$, the estimated radience from each input image. To compare the images, use the commands `imshow(log(E1),[-4 4])`, `imshow(log(E2),[-4 4])`, `imshow(log(E3),[-4 4])`. This plots each image on a log scale, and sets the ranges ($[-4, 4]$) to be the same. To what parts of the composite does each image contribute? Hint: you can also look at the "ownership" images for I_j.
3. What is the role of the smoothing term λ? To demonstrate its effect, try fitting the response curve with $\lambda = 10$ (low value) and $\lambda = 1000$ (high value). Print out the plots of the response functions and comment on your findings.

3. Lightness and shading [12 marks]

(a) Lambertian blocks [4 marks]

Suppose we observe a trihedral junction (ie., a vertex where three faces meet) of a rectangular block, with known orientation wrt the camera. If the block has a Lambertian surface of constant albedo, we can determine the direction of a single (distant) light source from the relative brightness of the three faces of the block.

If the orientation of each face in the camera's reference frame is given by unit normal vectors $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$, and $\hat{\mathbf{n}}_3$, respectively, show how to obtain the lighting direction and albedo of the surface.

Note: You can ignore the $\cos^4\theta$ term in the imaging equation and that we are using orthographic projection.

(b) Derivation of photometric stereo [3 marks]

(Adapted from Horn, Exercise 10-16) Suppose we have three images of a Lambertian surface under three different light source directions.

The three images E_1, E_2, and E_3 are given by the following equations,

$$E_1 = \rho(\hat{\mathbf{s}}_1 \cdot \hat{\mathbf{n}}), \quad E_2 = \rho(\hat{\mathbf{s}}_2 \cdot \hat{\mathbf{n}}), \quad E_3 = \rho(\hat{\mathbf{s}}_3 \cdot \hat{\mathbf{n}})$$

where ρ is the surface albedo (may vary with position (x, y)), $\hat{\mathbf{n}}$ is the surface normal (unit vector), and

$$\hat{\mathbf{s}}_i = \frac{(-p_i, -q_i, 1)^T}{\sqrt{1 + p_i^2 + q_i^2}}$$

for $i = 1, 2, 3$ give the directions of the light source.

By subtracting pairs of equations, show that:

$$\rho\hat{\mathbf{n}} \cdot (E_1\hat{\mathbf{s}}_2 - E_2\hat{\mathbf{s}}_1) = 0 \;\text{ and }\; \rho\hat{\mathbf{n}} \cdot (E_2\hat{\mathbf{s}}_3 - E_3\hat{\mathbf{s}}_2) = 0.$$

The above equation shows that $\hat{\mathbf{n}}$ is perpendicular to both $(E_1\hat{\mathbf{s}}_2 - E_2\hat{\mathbf{s}}_1)$ and $(E_2\hat{\mathbf{s}}_3 - E_3\hat{\mathbf{s}}_2)$. Show that $\hat{\mathbf{n}}$ must be parallel to

$$(E_1\hat{\mathbf{s}}_2 - E_2\hat{\mathbf{s}}_1) \times (E_2\hat{\mathbf{s}}_3 - E_3\hat{\mathbf{s}}_2) = E_2(E_1(\hat{\mathbf{s}}_2 \times \hat{\mathbf{s}}_3) + E_2(\hat{\mathbf{s}}_3 \times \hat{\mathbf{s}}_1) + E_3(\hat{\mathbf{s}}_1 \times \hat{\mathbf{s}}_2))$$

Conclude that

$$\rho\hat{\mathbf{n}} = k(E_1(\hat{\mathbf{s}}_2 \times \hat{\mathbf{s}}_3) + E_2(\hat{\mathbf{s}}_3 \times \hat{\mathbf{s}}_1) + E_3(\hat{\mathbf{s}}_1 \times \hat{\mathbf{s}}_2))$$

for some constant k.

By taking the dot product of this relationship with $\hat{\mathbf{s}}_1$, and remembering that $E_1 = \rho(\hat{\mathbf{s}}_1 \cdot \hat{\mathbf{n}})$, show that

$$\rho(\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}_1) = kE_1\left[\hat{\mathbf{s}}_1\hat{\mathbf{s}}_2\hat{\mathbf{s}}_3\right]$$

where $[\hat{\mathbf{s}}_1\hat{\mathbf{s}}_2\hat{\mathbf{s}}_3]$ is the triple product $\hat{\mathbf{s}}_1 \cdot (\hat{\mathbf{s}}_2 \times \hat{\mathbf{s}}_3)$.

Finally, show that $k = \frac{1}{[\hat{\mathbf{s}}_1\hat{\mathbf{s}}_2\hat{\mathbf{s}}_3]}$, so that

$$\rho\hat{\mathbf{n}} = \frac{(E_1(\hat{\mathbf{s}}_2 \times \hat{\mathbf{s}}_3) + E_2(\hat{\mathbf{s}}_3 \times \hat{\mathbf{s}}_1) + E_3(\hat{\mathbf{s}}_1 \times \hat{\mathbf{s}}_2))}{[\hat{\mathbf{s}}_1\hat{\mathbf{s}}_2\hat{\mathbf{s}}_3]}$$

(c) Photometric stereo experiment [5 marks]

I have given you three images of an unknown surface illuminated from three lightsource directions: $\mathbf{s}_1 = [0, 0, 1]^T, \mathbf{s}_2 = [-1, 1, 2]^T, \mathbf{s}_3 = [1, 0, 3]^T$. Assuming the light source has the same radiance at each direction (ie., $E_1 = E_2 = E_3$), use the expression in part (b) to compute the surface normal and albedo at each point. To accomplish this, complete the program `photometric.m`, and show plots of surface normals and albedo ρ. Note: Since we don't know E, we can only determine ρ only up to a constant factor. *Remember to normalize the light source directions* $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$ *to unit vectors in your code.*