Homework #1

Displays, Alpha Compositing, Image Processing,

Affine Transformations, Hierarchical Modeling

Assigned: Friday, April 12th

Due: Thursday, April 25th
at the beginning of class

Directions: Please provide short written answers to the following questions on your own paper. Feel free to discuss the problems with classmates, but please ***answer the questions on your own and show your work.***

Please write your name on your assignment!

Problem 1: Short answer (10 points)

a) (6 Points) Much of our perception of 3D arises from the fact that we have two eyes, viewing a scene from two different viewpoints, so-called stereo vision. Given images recorded or rendered for two different viewpoints separated by the typical distance between human eyes, a 3D stereo display presents one image to one eye and the other image to the other eye; our eyes are then fooled into believing they are looking at an actual 3D scene. One approach to 3D displays is based on a standard color LCD display and a pair of LCD shutter glasses. The LCD display first shows a left-eye image, then a right-eye image, and so on, while the LCD shutter glasses synchronously let light reach the left eye, then the right, etc. An LCD shutter is essentially one giant LCD pixel (a crystal sandwiched between two polarizers) with no color filter, driven with a voltage to be either opaque or transmissive. The display and shutters are designed to give a reasonably bright picture when sitting naturally in front of the display. Assume that the LCD crystal at each display pixel is oriented the same way as every other pixel in the display (regardless of color filter).

- If you tilt your head sideways (i.e., tilting your head over to one of your shoulders, so that the imaginary line segment connecting your eyes is now aligned with the vertical direction), will the displayed images appear dimmer in one eye, both eyes, or neither eye? Justify your answer.

- Suppose you removed the LCD filter panel in front of the unpolarized backlight and looked at the even backlighting with your naked eye(s); you would see even, white light. Roughly how much dimmer would you expect that light to become after putting the panel back on and putting on the shutter glasses (which are turned on and shuttering), assuming the framebuffer is set to white at each pixel for both eyes? [Note that unpolarized light intensity is cut in half by linear polarization. Assume that the R,G,B sub-pixels each transmit 1/3 of the visible spectrum of the light. Perceived brightness is averaged over time.] Justify your answer.

b) (4 Points) Consider two vectors u and v which are of non-zero length and not parallel to each other. Which of the following is true and which is false:

$$(u \times v) \times u = u \times (v \times u)$$

$$\left[(u \times v) \times u\right] \cdot u = 0$$

$$\left[(u \times v) \times u\right] \cdot v = 0$$

$$\left\{ \frac{u}{\|u\|} \times \left[\frac{(u \times v) \times u}{\|(u \times v) \times u\|} \right] \right\} \cdot \left\{ \frac{u}{\|u\|} \times \left[\frac{(u \times v) \times u}{\|(u \times v) \times u\|} \right] \right\} = 1$$

You do **not** need to justify your answer.

Problem 2: Alpha compositing (18 points)

The alpha channel is used to control blending between colors. The most common use of alpha is in "the compositing equation"

$$\mathbf{C} = \alpha\,\mathbf{F} + (1-\alpha)\,\mathbf{B} \quad \text{or} \quad \begin{bmatrix} C_R \\ C_G \\ C_B \end{bmatrix} = \alpha \begin{bmatrix} F_R \\ F_G \\ F_B \end{bmatrix} + (1-\alpha) \begin{bmatrix} B_R \\ B_G \\ B_B \end{bmatrix}$$

where α is the blending coefficient, **F** is the foreground color, **B** is the background color, and **C** is the composite color. In film production, compositing is a common operation for putting a foreground character into a new scene (background). The challenge faced with real imagery is to extract per pixel alpha and foreground color from a live action sequence, to enable compositing over a new background.

(a) (4 points) When filming an actor, a color **C** is observed at each pixel. If the three observed color channel values C_R, C_G, and C_B are the only knowns at a given pixel, how many unknowns remain in the compositing equation at that pixel? Treating each color channel separately, how many equations are there at the pixel? Is it generally possible to solve for all the unknowns under these circumstances? [Note: we are treating each pixel in isolation, so in each of these problems, you should just be thinking in terms of a single pixel.]

(b) (2 points) To assist the process of extracting the desired **F** and α values, the actor may be filmed against a known background, typically solid blue or green. If the components of **B** are known, how many unknowns remain at a given pixel? Is it possible, in general, to solve for **F** and α under these circumstances?

(c) (6 points) When filming the original Star Wars trilogy, the starships were assumed to contain only shades of gray and were filmed against a solid blue background. Thus, at a given pixel, the visual effects people could assume $\mathbf{F} = [L\ L\ L]^T$, where L is a shade of gray, and $\mathbf{B} = [0\ 0\ 1]^T$, where color channel values are in the range [0...1]. Given an observed color $\mathbf{C} = [C_R\ C_G\ C_B\,]^T$ at a pixel, compute α and L in terms of the color components of **C**. You should comment on how to handle the case when $\alpha = 0$. Show your work. [Note: if the answer is not unique, just provide one possible solution.]

(d) (6 points) Suppose you had the luxury of two consecutive images of a stationary foreground subject against a blue and a green background in succession, $\mathbf{B} = [0\ 0\ 1]^T$ and $\mathbf{G} = [0\ 1\ 0]^T$, thus recording two colors, **C** and **D**, respectively, at each pixel. You would then have to consider two color compositing equations $\mathbf{C} = \alpha\,\mathbf{F} + (1-\alpha)\,\mathbf{B}$ and $\mathbf{D} = \alpha\,\mathbf{F} + (1-\alpha)\,\mathbf{G}$. Solve for α and the components of the foreground color, F_R, F_G, and F_B at a given pixel in terms of the components of **C** and **D**. Show your work. [Note: if the answer is not unique, just provide one possible solution.]

Problem 3: Image Processing (16 points)

Suppose we have two filters:

0	0	0
-1	0	1
0	0	0

Filter *A*

0	1	0
0	2	0
0	1	0

Filter *B*

a) (3 points) In class, we described a simple and intuitive version of an *x*-gradient filter: [-1 1]. When applied, this filter computes the *finite difference* gradient in the *x*-direction, essentially solving for $\partial f / \partial x \approx \Delta f / \Delta x$, where $\Delta x = 1$ and pixels are one unit distance from their neighbors. Filter *A*, by contrast, is used to compute what is known as the *central difference* *x*-gradient. Although it cannot be normalized in the usual way, since its values sum to zero, it is usually multiplied by a coefficient of ½. Why?

b) (3 points) Normalize *B*. What effect will this normalized filter have when applied to an image?

c) (4 points) Compute *A*B*, using *A* and *B* from the ***original*** problem statement, i.e., ***without*** using the scale factors described in (a) and (b). You can treat *B* as the filter kernel and assume that *A* is zero outside of its support. You do *not* need to show your work. [Aside: convolution is commutative (*A*B=B*A*), so you would get the same answer by using *A* as the filter kernel. But, you would have to remember to "flip" the kernel to get $\widetilde{A}[i, j] = A[-i, -j]$. We've asked you instead to use *B* as the filter kernel, but since *B* is symmetric, i.e., $\widetilde{B}[i, j] = B[-i, -j] = B[i, j]$, you don't need to worry about flipping.]

d) (2 points) Compute *A*B*, now using *A* and *B* after scaling them according to (a) and (b).

e) (4 points) If we apply the result of (c) or (d) to an image *f*, we are computing (*A*B*)**f*. Convolution is associative, so we would get the same results as computing *A**(*B*f*). In other words, we're filtering the image with *B*, and then filtering the result with *A*.

- Why would it be desirable to apply *B* before computing the gradient (as opposed to not applying *B* at all)?

- Why might applying *B* be better than applying a filter *B*' that is filled with a full 3x3 set of positive coefficients (e.g., changing the 0's to 1's in the *B* filter given above), rather than just a single column of coefficients?

Problem 4: 3D Coordinate Systems (14 points)

Recall that the canonical coordinate system can be defined in terms of the set of canonical *x*-, *y*-, and *z*-axes and corresponding canonical vectors, $\hat{\mathbf{x}} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$, $\hat{\mathbf{y}} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$, and $\hat{\mathbf{z}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$. (Here we ignore the homogenous component and assume we are working with linear, not affine, coordinates and transformations.)

Suppose we perform a transformation to a new coordinate system, which maps the canonical axis vectors to new axis vectors ($\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{w}}$). I.e., the transformation maps $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ to $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{w}}$, respectively. We define $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{w}}$ to have similar properties to the canonical vectors: they are each of unit length and orthogonal (perpendicular) to each other, and they follow the right-hand rule, i.e., if you curl the fingers of your right hand from $\hat{\mathbf{u}}$ to $\hat{\mathbf{v}}$ your thumb will point in the direction of $\hat{\mathbf{w}}$.

All of the canonical vectors are of unit length and orthogonal (perpendicular) to each other.

a) (3 points) Suppose we form a 3x3 matrix, each column filled with a canonical vector:

$$M = \begin{bmatrix} \hat{\mathbf{u}} & \hat{\mathbf{v}} & \hat{\mathbf{w}} \end{bmatrix}$$

What kind of geometric transformation does this describe? Explain.

b) (5 points) Solve for $M^T M$, simplifying as much as possible. You can solve in terms of $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{w}}$ without expanding them into their components. Or you can explicitly form matrices after expanding the vectors into $\hat{\mathbf{u}} = \begin{bmatrix} u_x & u_y & u_z \end{bmatrix}^T$, etc., to fill the matrices. Show your work. What does your answer imply about M^T?

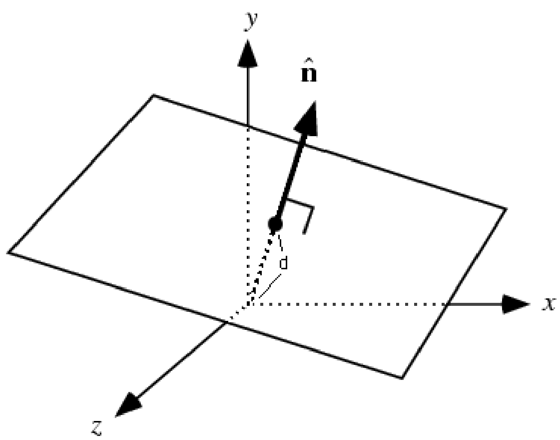
c) (6 points) Now suppose you are given two vectors, **a** and **b**, each of arbitrary, non-zero length. We will also assume that these vectors are linearly independent, $\mathbf{a} \neq \alpha\, \mathbf{b}$ for any scalar value of α; i.e., the vectors don't point in exactly the same or exactly opposite directions. We can use these vectors to form a right-handed coordinate system, $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, $\hat{\mathbf{w}}$, where:

- $\hat{\mathbf{w}}$ points in the same direction as **b**
- $\hat{\mathbf{u}}$ is orthogonal to both **a** and **b**.

Solve for $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{w}}$, each in terms of **a** and **b**. You do not need to expand any vectors into their components; e.g., if you need to take a dot product between **a** and **b**, you can just refer to $\mathbf{a} \bullet \mathbf{b}$ without expanding it into the sum of the products of components. Note that there is more than one possible answer to this problem; you should provide just one solution.

Problem 5: 3D Affine Transformations (18 points)

The equation $\hat{\mathbf{n}} \bullet \bar{\mathbf{x}} = d$ describes the plane pictured below which has unit length normal $\hat{\mathbf{n}}$ pointing away from the origin and is a distance d from the origin (in the direction of the normal vector). Any point $\bar{\mathbf{x}} = [x \quad y \quad z]$ on the plane must satisfy the plane equation $\hat{\mathbf{n}} \bullet \bar{\mathbf{x}} = d$.

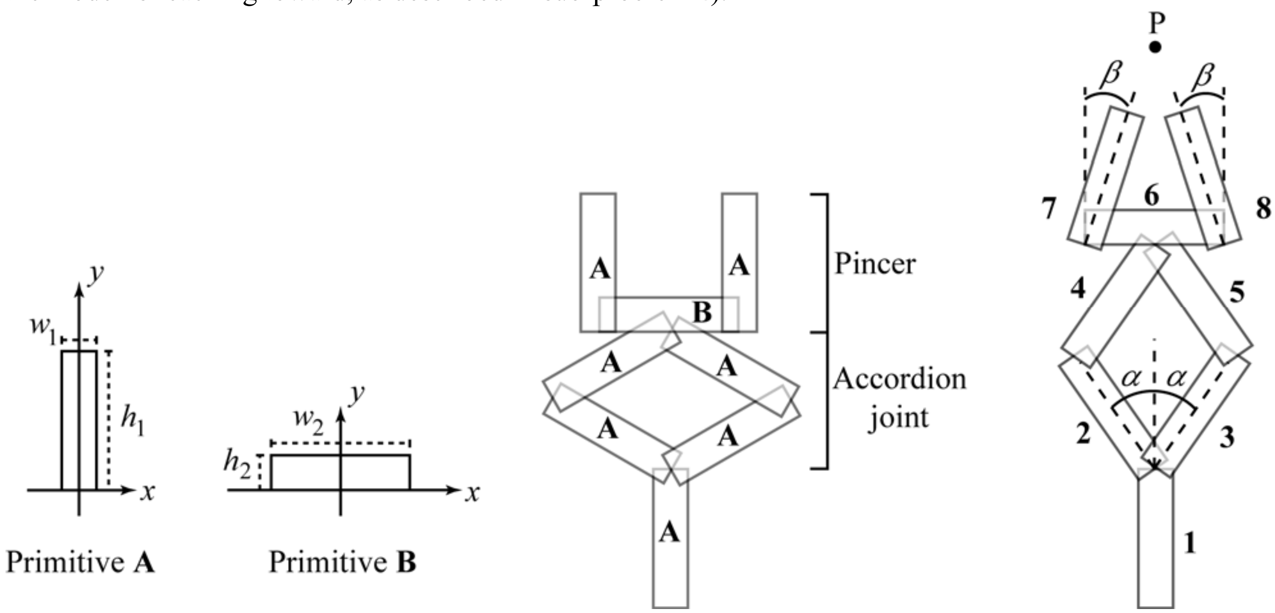


$$M_{xy} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Now consider a plane with normal lying in the *y*-*z* plane. The normal will have the form (0, sinθ, cosθ) for some θ. The equation for the plane is then $y\sin\theta + z\cos\theta = d$. Write out the product of 4x4 matrices that would perform a reflection across this plane. One of these matrices will be a reflection matrix; you must use the matrix M_{xy} above, which performs a reflection across the *x*-*y* plane. You must write out the elements of the matrices and the product order in which they would be applied, but you do not need to multiply them out. Justify your answer with words and/or drawings.

Problem 6: Hierarchical modeling (24 points)

Suppose you want to model the pincer with accordion joint illustrated below. The model is comprised of 8 parts, using primitives **A** and **B**. The model is shown in two poses below, with the controlling parameters of the model illustrated on the far right. The illustration on the right also shows a point P that the model is reaching toward, as described in sub-problem **c**).



Assume that α and β can take values in the range [0, 90°]. Also assume that all parts use primitive **A**, except for part **6**, which uses primitive **B**. The model on the left shows the primitives used, the model on the right shows the enumeration (naming) of the parts.

The following transformations are available to you:

- R(θ) – rotate by θ degrees (counter clockwise)
- T(a, b) – translate by $\begin{bmatrix} a \\ b \end{bmatrix}$

a) (16 points) Construct a tree to describe this hierarchical model using part **1** as the root. Along each of the edges of the tree, write expressions for the transformations that are applied along that edge, using the notation given above (you do not need to write out the matrices). Remember that the order of transformations is important! Show your work wherever the transformations are not "obvious." Your tree should contain a bunch of boxes (or circles) each containing one part number (1…8); these boxes should be connected by line segments, each labeled with a corresponding transformation that connects child to parent. The tree must have one or more branches in it. If two parts are connected physically, then they should be connected in the tree, as long as you don't form a cycle by connecting them.

b) (2 points) Write out the full transformation expression for part **7**.

c) (6 points) Suppose the primitives are infinitesimally thin, $w_1 = h_2 = 0$, and have lengths $h_1 = 10$ and $w_2 = 12$. Assume that part **1** sits right on the origin in world coordinates. What would the α and β parameters have to be so that the model extends out and closes the pincer just enough to precisely grasp the point $P = \begin{bmatrix} 0 & 28 \end{bmatrix}^T$, in world coordinates. Show your work.